

DDB	361	QYVERKLPIMRGSPVVDKNTYAFRGFLDKYAEDSSSTIKWIDAPDCFCFHWNAAEPE	4.20
QY	4.21	TDEVVIGSCMPPDSDNESENLSKSVLSERILNLKGESETRPTISNDQYVNLQAGM	4.80
Db	4.21	TDEVVIGSCMTPDSDNESENLSKSVLSERILNLKGESETRPTISNDQYVNLQAGM	4.80
QY	4.81	VNRNMLGRKTKEFAYLALAEPWPVKSGFAKVDLTIGEVKKHLYGDNRYYGEPFLPGEGE	5.40
Db	4.81	VNRNMLGRKTKEFAYLALAEPWPVKSGFAKVDLTIGEVKKHLYGDNRYYGEPFLPGEGE	5.40
QY	541	EDEGYILCQYHDEKTKWKSPLQIVNAVSLEVEATYKLPSRVYPQFGHGTIGADDLKQWV	5.99
Db	541	EDEGYILCQYHDEKTKWKSPLQIVNAVSLEVEATYKLPSRVYPQFGHGTIGADDLKQWV	5.99
RESULT 3			
	AAE04789	AAE04789 standard; protein: 605 AA.	
	XX		
	AC	AAE04789;	
	XX		
	DT	10-SEP-2001 (first entry)	
	XX		
	DE	Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1.	
	XX		
	XX	Tomato; neoxanthin cleavage enzyme; LeNCED1; abscisic acid; ABA; stress tolerance; transgenic plant; antisense; antisense-therapay; plant breeding; herbicide.	
	XX		
	OS	KW	
	XX	KW	
	OS	KW	
	XX	KW	
	OS	KW	
	XX	KW	
	OS	KW	
	XX	KW	
	PN	Lycopersicon esculentum.	
	XX		
	EP1116794-A2,		
	XX		
	PD	18-JUL-2001.	
	XX		
	XX	11-JAN-2001; 2001EP-0300218.	
	XX		
	PR	13-JAN-2000; 2000JP-0010056.	
	PR		
	PR	11-JAN-2001; 2001JP-0003476.	
	XX		
	PA	(RIKE) RIKEN KK.	
	XX		
	PT	Iuchi S, Kobayashi M, Shinozaki K;	
	XX		
	WPI:	2001400081/43.	
	DR	N-PSDB; AAD09401.	
	XX		
	PT	A DNA encoding a protein with a neoxanthin cleavage activity for producing transgenic plants with improved or decreased stress tolerance	
	PT		
	PT		
	XX		
	PS	Claim 3; Fig 2; 101pp; English.	
	XX		
	CC	The invention relates to neoxanthin cleavage enzymes and their corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key role in endogenous abscisic acid (ABA) biosynthesis under drought stress.	
	CC	Neoxanthin cleavage enzyme is used for improving stress tolerance in a plant when expressed in a plant cell. The invention also relates to methods for increasing or decreasing stress tolerance in a plant by introducing the DNA into the plant, and a transgenic plant into which neoxanthin cleavage enzyme is introduced. The improvement of stress tolerance in plants is useful, for example, in plant breeding. Neoxanthin cleavage enzyme genes are useful for producing transgenic plants. An a land can be improved by growing transformatant weed for several years and then removing the weed by specifically lowering stress tolerance in the weed by introducing an inducible promoter. The present sequence is Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1 protein related to the invention.	
	XX		
	Sequence	605 AA;	
	SQ	72.4%; Score 2280.5; DB 22; Length 605;	

PA (PION-) PIONEER HI-BRED INT INC. (CURA-) CURAGEN CORP.	DB 478 QVNLEAGMVRNKLGRKTEVAYLATEAPWPKVSFGEVKLVNLTFGEVERKFIVCDNKYGGEP.
XX	QY 533 FLPGEG-GGEEDEGYLLCFVHDEKWKSELQIVNAVSLEAVATKPLSPRVYGEFHSTFIG 590
XX	XX
PT	Db 538 FLPRDNNSEEDGYLLAFVHDEKWKSELQIVNAVSLEAVATKPLSPRVYGFHTFIN 597
PP Novel isolated defence-related signalling gene isolated from sunflower protein encoding neoxanthin cleavage enzyme, amino acid permease or glutamic acid-rich protein useful for increasing resistance of plant to a pathogen	QY 591 ADDLAKQ 597
PP	Db 598 ANDLANQ 604
XX	
PS Disclosure; Fig 1; 135pp; English.	RESULT 5
XX	AAE04787 standard; Protein: 612 AA.
CC This invention relates to defence-related signalling genes isolated from the sunflower (<i>Helianthus annuus</i>). The genes encode a neoxanthin cleavage enzyme (NCE), an amino acid permease (AAP) and a glutamic acid rich, protein (GRP). The signalling gene is useful for increasing the resistance of a plant to a pathogen such as fungus, virus, bacterium, nematode or insect (e.g. European corn borer), preferably Sclerotinia spp., <i>Phoma spp.</i> , or <i>Phomopsis spp.</i> , by stably incorporating a construct containing the gene into the genome of the plant. The gene is useful for regulating gene expression in a plant, in response to a stimulus such as infection with a pathogen, damage from a pathogen, hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid, oxalic acid or expression of a gene encoding oxalate acid oxidase. The genes are also useful for stem-preferred regulation of gene expression in a plant. The genes are useful in agriculture, particularly in the breeding of crop plants with improved agronomic traits, for modifying abscisic acid (ABA) metabolism and for modifying amino acid transport and content in plants. The present sequence represents a neoxanthin cleavage enzyme-like protein from <i>Lycopersicon esculentum</i> used in the characterisation of sunflower NCE.	ID AAE04787 XX AC AAE04787; XX DT 10-SEP-2001 (first entry)
CC Sequence : 605 AA;	DE Vigna unguiculata neoxanthin cleavage enzyme, CPRD65, XX Cowpea; neoxanthin cleavage enzyme; abscisic acid; ABA; herbicide; stress tolerance; transgenic plant; plant breeding; antisense-therapy; plant growth protectant; Cowpea Responsive to Dehydration; CPRD65. XX OS Vigna unguiculata. XX EP116794-A2; XX PR 13-JAN-2000; 20000JP-0010056. PR 11-JAN-2001; 2001JP-0003476. XX (RIKE) RIKEN KK.
CC	PJ Iuchi S, Kobayashi M, Shinozaki K; XX DR WPI; 2001-460081/43. N-BSDB; AAD0399.
CC	XX PT A DNA encoding a protein with a neoxanthin cleavage activity for producing transgenic plants with improved or decreased stress tolerance XX PR 11-JAN-2001; 2001JP-0003476. PR 13-JAN-2000; 20000JP-0010056. XX PS Claim 3; Fig 2; 101pp; English.
CC	The invention relates to neoxanthin cleavage enzymes and their corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key role in endogenous abscisic acid (ABA) biosynthesis under drought stress. Neoxanthin cleavage enzyme is used for improving stress tolerance in a plant when expressed in plant cell. The invention also relates to methods for increasing or decreasing stress tolerance in a plant by introducing the DNA into the plant, and a transgenic plant into which a neoxanthin cleavage enzyme is introduced. The improvement of stress tolerance in plants is useful, for example in plant breeding. Neoxanthin cleavage enzyme genes are useful for producing transgenic plants. An arid land can be improved by growing transgenic weed for several years and then removing the weed by specifically lowering stress tolerance in the weed by inducing an inducible promoter. The present sequence is Vigna unguiculata neoxanthin cleavage enzyme, CPRD65 (Cowpea Responsive to Dehydration) protein. CPRD65 gene is isolated from cowpea plant. XX Sequence 612 AA;
CC	Query Match 72.4%; Score 2380.5; DB 22; Length 605; Best Local Similarity 71.7%; Pred. No. 5.3e-222; Matches 435; Conservative 63; Mismatches 90; Indels 19; Gaps 6;
QY 5 TATAAVSGRWLGGNHTOPPLSSQQSSLQSYCS--SLPMAASRVTRKLNVSSALHTPPALH 61	Db 3 TTTSHTATNTW----KTKLSPDSSKKEFQASNSISLQNINRQSNINSLQAPPALH 57
Db 62 FPKQSSN--SPA--IVKPKAKAESN----TKOMNLFORAAALDAEAEGLYSHEKLHP 112	QY 58 FPKQSSNQTPKNTISPKQNNNNSSSSSKWNLYQKAAMALDAVEASLTKHELEHP 117
Db 58 FPKQSSNQTPKNTISPKQNNNNSSSSSKWNLYQKAAMALDAVEASLTKHELEHP 117	Db 113 LPKTADPSVQIAGNFAPNEQPYRNPLPVGKLPDSIKGVYVNRNGANDLHEPVTHFFD 172
QY 113 LPKTADPSVQIAGNFAPNEQPYRNPLPVGKLPDSIKGVYVNRNGANDLHEPVTHFFD 172	Db 118 LPKTADPSVQIISNNFAPENPYCQSLPVGKLPKCVOGVYVNRNGANDLHEPVTHFFD 177
Db 118 LPKTADPSVQIISNNFAPENPYCQSLPVGKLPKCVOGVYVNRNGANDLHEPVTHFFD 177	QY 173 GDDKMHAVKFEHQGSASTACRTFVINRQERQGRPVPKAIGELHGTGIAIRLMFYAR 232
QY 173 GDDKMHAVKFEHQGSASTACRTFVINRQERQGRPVPKAIGELHGTGIAIRLMFYAR 232	Db 178 GDDKMHAVQEKNSASTACRFVTERLYQEKALGRPVPKAIGELHGTGIAIRLMFYAR 237
Db 178 GDDKMHAVQEKNSASTACRFVTERLYQEKALGRPVPKAIGELHGTGIAIRLMFYAR 237	QY 233 AAGTVDPAHGTVGANAGLVYNGRLAMSSEDDLPYQVQTNGDLKTVGRFDGQLES 292
QY 233 AAGTVDPAHGTVGANAGLVYNGRLAMSSEDDLPYQVQTNGDLKTVGRFDGQLES 292	Db 238 GLGLDVHDISKGTGVANALVYFNRLAMSDLPYHVKVTGDLKIEGRDDEGQIKRS 297
Db 238 GLGLDVHDISKGTGVANALVYFNRLAMSDLPYHVKVTGDLKIEGRDDEGQIKRS 297	QY 293 TMTAHPKIDPESEGELFALSVDYNSKPKLYKFRSPDGTKSPDYEIQLDQPTNMHDFATE 352
QY 293 TMTAHPKIDPESEGELFALSVDYNSKPKLYKFRSPDGTKSPDYEIQLDQPTNMHDFATE 352	Db 298 TMTAHPKDPVSPGELFALSVDYQKPKLYKFRSPSKNSNDVEIPDPTNMHDFATE 357
Db 298 TMTAHPKDPVSPGELFALSVDYQKPKLYKFRSPSKNSNDVEIPDPTNMHDFATE 357	QY 353 NFFFVPPDQGVVFKLPEMIRGSPVWYDKNKVARGFLDKYAEDSSNIKWIDAPDCFCFHL 412
QY 353 NFFFVPPDQGVVFKLPEMIRGSPVWYDKNKVARGFLDKYAEDSSNIKWIDAPDCFCFHL 412	Db 358 NFFFVPPDQGVVFKLPEMIRGSPVWYDKNKVARGFLDKYAEDSSNIKWIDAPDCFCFHL 417
Db 358 NFFFVPPDQGVVFKLPEMIRGSPVWYDKNKVARGFLDKYAEDSSNIKWIDAPDCFCFHL 417	QY 413 QVNLEAGPERDEVVYIGSCMTPDPSDNESENLSNLKGTGESTRPLISNEDQ 472
QY 413 QVNLEAGPERDEVVYIGSCMTPDPSDNESENLSNLKGTGESTRPLISNEDQ 472	Db 418 WNAMEEAETDIEVIVIGSCMTPDPSIFNEDEGLKSVSLIRLAKTGKSTRLIENDE 477
Db 418 WNAMEEAETDIEVIVIGSCMTPDPSIFNEDEGLKSVSLIRLAKTGKSTRLIENDE 477	QY 473 QVNLEAGMVNRMLGRKTKFAYLALAEPWPWKVSFGAKVDTLTIGEVKKHLYGDNRYGGEP.
QY 473 QVNLEAGMVNRMLGRKTKFAYLALAEPWPWKVSFGAKVDTLTIGEVKKHLYGDNRYGGEP.	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-
	Score 2168.5; DB 22; Length 612;
	Best Local Similarity 69.6%; Pred. No. 1.3e-210;
	Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;
	Sequence 612 AA;
	Query 8 AAVSGRWLGGNHTQPPLSSQSSDLSYCLPMASRYTRKLN-VSSAIIHTPPALHFHKQ-</td

Db 3 SSASNTWNATLSPSPEKKDPLSTS-SPTNQIPL--RKTSSNTITCSLQI---LHEPKQY 56
 Qy 66 ---SSNSPATIVVKPKAKESNH-----KONNLFRQAAAALDA 100
 Db 57 QPTSTESTATTTTPPKTTTTPPKTTTTPPRETNPLSDTNQPLQKWNFLOKAAAPALDV 116
 Qy 101 EGFLVSHKEKLHPLPKTADPSYQSIAGNFAPIQPYRNQVGRVNGANP 160
 Db 117 ETALVSHERKFPLPKTADPSYQSIAGNFAPIQPYRNQVGRVNGANP 176
 Qy 161 LHEPVYGHFFPDGDMVHAYKFEHGSSAYACRFQTQNRFQERQLGRPVPKAIGELGH 220
 Db 177 LYEPVAGHFFPDGDMVHAYKETNGASAYACRFETQRLSOKSLGPVPKAIGELGH 236
 Qy 221 TGIARMLFYARAAGIVDPAHGTGYANAGLVFNGRLLAMSEDDLPYQVQITPNSDLKT 280
 Db 237 SGSTARLILFYARGLFLVNDQGMGVANAGLVFNFNNHLLAMSEDDLPYHWRITPNSDLT 296
 Qy 281 YGRFDEDFGQLEISTMIAHPKVDPGKTDPESELFA-SYDVSKPYKFRISPDKGKSPDVE-QLD 340
 Db 297 VGRYDFNGOLNSTMIAHPKLDPPVGDLHALSYDVQKPYKFRISPDKGKSPDVEIPLK 356
 Qy 341 QPTMMHDFAITENFVWVDPQQVVFKLPEMTRGSPVWYDKNKVAPFGILDKYAEDSSNIK 400
 Db 357 EPTMMHDFAITENFVWVDPQQVVFKLTEMIGSPVWYDKNKTSREGILHNKAADANAMR 416
 Qy 401 WIDAPDCFCFLWNAWEPEETDVYTGSCMTPPDSIFENEDENIKSVLSSIRLANKTG 460
 Db 417 WIDAPDCFCFLWNAWEPEETDVYTGSCMTPADSFENEBEUSIKSVLSSIRLNRGK 476
 Qy 461 STRRPLISNEEQVNLEAGMVNRNMGLRKTFAYALAEPPKVSFAKYLGEVKK 520
 Db 477 STRRPLIS-DAEQVNLEAGMVNRNMGLRKTOFAYALAEPPKVSFAKYLGEVKK 535
 Qy 521 LYGDNYRGEPFPLPPEGGEDEGYTLCPVIDEKTKWSELOVNNAVSLEYEVTKLPSRV 580
 Db 536 MYGEERFGEPEFLP-NGOKEEDGTYLAVHDEKEWKSELOVNNAQNLKEASIKLPSRV 594
 Qy 581 PYGFHOTFIGADDLAKQ 597
 Db 595 PYGFHOTFIGADDLAKQ 611

RESULT 6
 ID ABB91182 standard; protein; 589 AA.
 XX
 AC ABB91182;
 XX
 DT 31-MAY-2002 (first entry)
 XX
 DE Herbicidally active polypeptide SEQ ID NO 393.
 XX
 KW Herbicidal; plant; agriculture; herbicide.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2002102010-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 28-AUG-2001; 20001WO-EPO9892.
 XX
 PR 28-AUG-2001; 2001WO-EPO9892.
 XX
 PA (PARB) BAYER AG.
 XX
 PI Tietjen K, Weidler M;
 XX
 DR WPI; 2002-269010/31.
 XX
 PT Identifying plant target proteins for herbicidally active compounds,

comprising aligning and comparing nucleic acid or amino acid sequences from non-plant
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 XX
 PS SEQ ID NO 393; 261pp + Sequence Listing; English.
 XX
 CC The invention relates to identifying target Proteins
 CC comprising herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC used as herbicides.
 XX
 Sequence 589 AA;
 Query Match 64.9%; Score 2044.5; DB 23; Length 589;
 Best Local Similarity 68.1%; Pred. No. 4.7e-18;
 Matches 390; Conservative 69; Mismatches 101; Indels 13; Gaps 4;
 Qy 27 SOSSDLSYCSSLPMASRVTKLNSSALHTPFPKQSSNSPAIVVKPKAKESTIKQM 86
 Db 27 SPSSSYSETNPKPR--RRKLSANSVSDTPALLNPNNPSPNPNI----PEKDTSRW 76
 Qy 87 NLFORAAAAALDAAEGLTFLVSHKHLPLKTAEDPSQVQAGNAPVNPQVVRNLPVVGKLP 146
 Db 77 NPLQRASAALDEAFLRERSKPLPKTVQDPSVKSLSVSGKIP 136
 Qy 147 DSIKGYYVRNGANPLHEPTGHHFEDGDMVHAYKREHGSASYACRFQTQTRREVQRQLG 206
 Db 137 DCIDGFLYLRNGANPLHEPTGHHFEDGDMVHAYKREHGSASYACRFQTQTRREVQRQLG 196
 Qy 207 RPVFPKAIGELEGHHTGILFYARAAGIVDPAIGTGYANAGLYFNGRLLAMBEDDL 266
 Db 197 SPIFPKAIGELEGHSGIARMLFYARGLGFLGLINHKNGTGYANAGLYFHDRLLAMESDDL 256
 Qy 267 PYQVQITPQDGLIKTQGRFDQGQLESTMIAHPKVDPESGEFLFALSYDVSKPYKLYFRFS 326
 Db 257 PYQVRIVDQGDLTIEFLGQPLASPMHMDETAENEVVLPDQVQVKLMSFLRSVYDGEITSRF 316
 Qy 327 PGDTKSPDVEIOLDQTMHMDETAENEVVLPDQVQVKLMSFLRSVYDGEITSRF 376
 Db 317 PEGEKSPDVEIOLDQTMHMDETAENEVVLPDQVQVKLMSFLRSVYDGEITSRF 376
 Qy 387 GILDKYREDSSNIKWDADPDCFCFLWNAWEPEETDVVIGSCMPDSDIFENEDENLK 446
 Db 377 GTLPNRAKDASEMWVWSPETPCFHLMWNAWEPEETDVVIGSCMPDSDIFENCEQLN 436
 Qy 447 SVLSEIILNLNTKGKSTARLTIPG-SV7QMNLEAGMVNRNLGRKTKEYLAKAEPWPKVSG 506
 Db 437 FAKVDLUTGEVKHLYGDNRYGEPFLPQG-EGGEDEGGYTLCPVHDEKTWKSELOVN 564
 Qy 507 FAKVDLUTGEVKHLYGDNRYGEPFLPQG-EGGEDEGGYTLCPVHDEKTWKSELOVN 564
 Db 496 FAKVDLUTGEVKHLYGDNRYGEPFLPQG-EGGEDEGGYTLCPVHDEKTWKSELOVN 555
 Qy 565 AVSLEYEVATVKLPSRVYPYGFHGTFIGADDLAKQ 597
 Db 556 AVTLELFATVKLPSRVYPYGFHGTFVNSADMLNQ 588
 RESULT 7
 AAE04782
 ID AAE04782 standard; protein; 583 AA.
 XX
 AC AAE04782;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Arabidopsis thaliana neoxanthin cleavage enzyme, ANCEDI.

Neoxanthin cleavage enzyme; AtNCED1; abscisic acid; ABA; herbicide; stress tolerance; transgenic plant; plant breeding; antisense-therapy; plant growth protectant.

Arabidopsis thaliana.

EP1116794-A2.

PD 18-JUL-2001.

XX PF 11-JAN-2001; 2001EP-0300218.

XX PR 13-JAN-2000; 2000JP-0010056.

PR 11-JAN-2001; 2001JP-0003476.

XX PA (RIKE) RIKEN KK.

XX PI Iuchi S, Kobayashi M, Shinozaki K;

WPI: 2001-400081/43.

DR N-PSDB; AAD09394.

XX PR A DNA encoding a protein with a neoxanthin cleavage activity for producing transgenic plants with improved or decreased stress tolerance

PT -

XX PS; XX

Claim 3; Fig 10; 10pp; English.

The invention relates to neoxanthin cleavage enzymes and their corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key role in endogenous abscisic acid (ABA) biosynthesis under drought stress. Neoxanthin cleavage enzyme is used for improving stress tolerance in a plant when expressed in a plant cell. The invention also relates to methods for increasing or decreasing stress tolerance in a plant by introducing the DNA into the plant, and a transgenic plant into which a neoxanthin cleavage enzyme is introduced. The improvement of stress tolerance in plants is useful, for example in plant breeding. Neoxanthin cleavage enzyme genes are useful for producing transgenic plants. An arid land can be improved by growing transformatant weed for several years and then removing the weed by specifically lowering stress tolerance in the weed by inducing an inducible promoter. The present sequence is Arabidopsis thaliana neoxanthin cleavage enzyme, AtNED1 protein.

The AtNCED1 cDNA is obtained from an Arabidopsis plant-derived cDNA library using a cDNA of the CPRD65 (Cowpea Responsive to Dehydration) gene isolated from cowpea plant as a probe.

XX Sequence 583 AA;

Query Match 63 2%; Score 1991; DB 22; Length 583;

Best_Local_Similarity 64.5%; Pred. No. 1.2e-192;

Matches 380; Conservative 82; Mismatches 105; Indels 22; Gaps 6;

Db 16 GGAHTTOPPLSSQSSDLSY-CSSLPMASRVTKLNVSSALTPPAHFHKSSNSPAIV 73

Qy 17 GGKTWTW---QAQIDGSPRKRQPKVTCVQIDVTE-LTKKROLFTPTTAT?P--- 62

Db 74 VPKPAKESNTKQMNLFQRAAAALDAEGLFVSKHLPLKHDAPSQYIAGNFAPVNEQ 133

Qy 63 -----QINPLRNIFQAAAIAADAARALIHEQDSLPLKHDAPSQYIAGNFAPVNEQ 116

Db 134 PVRNLPLPVGKLPDSIKGVYVNRNGANPLHPEVGHFFDGDGMVHAYKFEGHSASYACRF 193

Db 117 SVRENLTIVEGTIPDCIDGVYVNRNGANPMEPTAGHFLDFGDKHVAKRITNGSASYACRF 176

Qy 194 TQFNRFVQERQLGRPFKAIGELHGHTGIARLMFLYARAAGIVDPAHGTGANAGLVY 253

Db 177 TKTERLYZEKRLGRPFKAIGELHGHTGIARLMFLYARGCLINNOQGVANAGLVY 236

Qy 254 GPGLLAMSESDLPLPQVQIPTNGDLKTVGREFDGOLESTMAHHPKVDESEGFALSYD 313

Db 237 FNNRLAMSEDLPYQLKITQTDLQTYGRDQKLPSAMAHPKLDPVTEHLHASYD_295

Qy 314 VWSKPLKYFRFSPDGTKSPDVEIQLDQOPTMMHDFAITTENFVVPDQOVVFKLPEMTRGG 373

Db 297 VVKKBYLKYFRFSPDGKSPPELEPLETPIMIHFAITENFVVPDQOVVFKLPEMISRK 356

Qy 374 SPVYDKNKVAREFGLDKYADSSNIKWDAPDCPCFHILWNAWBEPETDEVVIGSCMP 433

Db 357 SPVYFDGEYVSRLGIMPKAATEASOQIWINNSPEFCEHLWNAWBESPETEVVIGSCMP 416

Qy 434 PDSTFNESDENLKSVYLSETRLNLTGSTERRPILISNEDQVNLEAGMNRNMUGRKTKPFA 493

Db 417 ADSFNERDELSRSVLESERLNLTGSTERRPILISNEDQVNLEAGMNRNMUGRKTKPFA 474

Qy 494 YLAIAEPWPWKVSGFAKVDLTLTGEYKKHLYGDNRYGGPEFLPCEGG----EDEGYLICF 549

Db 475 FLAIAEPWPWKVSGFAKVDLCTGEMRKYYGKEKGPNNGENEEDGYPICFCH 534

Qy 550 VHDEKTWKSELQIYNAVASLEVEATVKKLPSRVPGFHGTIGADDLAKQY 598

Db 535 VHDEETKTSPLQIYNAVNLKLEATKKLPSRVPGFHGTIVFDSNLVDQL 583

RESULT 8

AAE04788

ID AAE04788 standard; Protein; 604 AA.

XX

AC AAE04788;

XX DT 10-SEP-2001 (first entry)

XX DE Zea mays neoxanthin cleavage enzyme, VP14.

XX KW Maize; neoxanthin cleavage enzyme; VP14; abscisic acid; ABA; herbicide; stress tolerance; transgenic plant; plant breeding; antisense-therapy; plant growth protectant.

XX KW

XX OS Zea mays.

XX PN EP1116794-A2.

XX PD 18-JUL-2001.

XX PF 11-JAN-2001; 2001EP-0300218.

XX PR 13-JAN-2000; 2000JP-0010056.

PR 11-JAN-2001; 2001JP-0003476.

XX PA (RIKE) RIKEN KK.

XX PI Iuchi S, Kobayashi M, Shinozaki K;

XX DR WPI: 2001-400081/43.

DR N-PSDB; AAD09400.

XX PT A DNA encoding a protein with a neoxanthin cleavage activity for producing transgenic plants with improved or decreased stress tolerance

XX PT

XX PS Claim 3; Fig 2; 10pp; English.

XX PA (RIKE) RIKEN KK.

XX PI Iuchi S, Kobayashi M, Shinozaki K;

XX DR WPI: 2001-400081/43.

DR N-PSDB; AAD09400.

XX PT A DNA encoding a protein with a neoxanthin cleavage activity for producing transgenic plants with improved or decreased stress tolerance

XX PT

XX PS Claim 3; Fig 2; 10pp; English.

XX PA (RIKE) RIKEN KK.

CC The invention relates to neoxanthin cleavage enzymes and their corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key role in endogenous abscisic acid (ABA) biosynthesis under drought stress.

CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a plant when expressed in a plant cell. The invention also relates to methods for increasing or decreasing stress tolerance in a plant by introducing the DNA into the plant, and a transgenic plant into which a neoxanthin cleavage enzyme is introduced. The improvement of stress tolerance in plants is useful, for example in plant breeding. Neoxanthin cleavage enzyme genes are useful for producing transgenic plants. An arid land can be improved by growing transformatant weed for several years and then removing the weed by specifically lowering stress tolerance in the weed by inducing an inducible promoter. The present sequence is Arabidopsis thaliana neoxanthin cleavage enzyme, AtNED1 protein.

CC The AtNCED1 cDNA is obtained from an Arabidopsis plant-derived cDNA library using a cDNA of the CPRD65 (Cowpea Responsive to Dehydration) gene isolated from cowpea plant as a probe.

CC Sequence 583 AA;

CC Query Match 63 Local_Similarity 64.5%; Pred. No. 1.2e-192; Length 583;

CC Best_Local_Similarity 64.5%; Pred. No. 1.2e-192; Length 583;

CC Matches 380; Conservative 82; Mismatches 105; Indels 22; Gaps 6;

CC Db 16 GGAHTTOPPLSSQSSDLSY-CSSLPMASRVTKLNVSSALTPPAHFHKSSNSPAIV 73

CC Qy 17 GGKTWTW---QAQIDGSPRKRQPKVTCVQIDVTE-LTKKROLFTPTTAT?P--- 62

CC Db 74 VPKPAKESNTKQMNLFQRAAAALDAEGLFVSKHLPLKHDAPSQYIAGNFAPVNEQ 133

CC Qy 63 -----QINPLRNIFQAAAIAADAARALIHEQDSLPLKHDAPSQYIAGNFAPVNEQ 116

CC Db 134 PVRNLPLPVGKLPDSIKGVYVNRNGANPLHPEVGHFFDGDGMVHAYKFEGHSASYACRF 193

CC Db 117 SVRENLTIVEGTIPDCIDGVYVNRNGANPMEPTAGHFLDFGDKHVAKRITNGSASYACRF 176

CC Qy 194 TQFNRFVQERQLGRPFKAIGELHGHTGIARLMFLYARAAGIVDPAHGTGANAGLVY 253

CC Db 177 TKTERLYZEKRLGRPFKAIGELHGHTGIARLMFLYARGCLINNOQGVANAGLVY 236

CC Qy 254 GPGLLAMSESDLPLPQVQIPTNGDLKTVGREFDGOLESTMAHHPKVDESEGFALSYD 313

CC Db 237 FNNRLAMSEDLPYQLKITQTDLQTYGRDQKLPSAMAHPKLDPVTEHLHASYD_295

CC Qy 314 VWSKPLKYFRFSPDGTKSPDVEIQLDQOPTMMHDFAITTENFVVPDQOVVFKLPEMTRGG 373

SQ	Sequence	604 AA;	PA (CURA-) CURAGEN CORP.
	Query Match	61.3%; score 1930; DB 22; Length 604;	XX PA
	Best Local Similarity	63.1%; Pred. No. 2.1e-186;	XX PI
	Matches	375; Conservative 73; Mismatches 122; Indels 22; Gaps 8;	XX Bidney DL, Crasta OR, Hu X, Lu G;
			XX WPI: 2001-211215/21.
QY	22	PPLSSQSSDLSYCSSLPMASRVTRKLNVSSALHTP---PALHFFPKQSSNSPATIVV--- 74	XX Novel isolated defence-related signalling gene isolated from sunflower
Db	17	PARSARAS---NSVRSPRAYSSVPPCETIOFFPKYADLPAPSRRKPAAPAVGHA 71	XX encoding neoxanthin cleavage enzyme, amino acid permease or glutamic
QY	75	-PKAKESNTQMNLEQRAAAAALDA-AEGFLVS-BERKLPPLPKTADPSVQIAGNEFAPVN 131	XX acid-rich protein useful for increasing resistance of plant to a
Db	72	AAPRKAEGGKQLNLFQRAAAAALDAEFSFVANLERSPFGLPLSTADPAVQIAGNFAPVG 131	XX pathogen -
QY	132	EQPVRNRLPVVGKLDPSIKGVYVRRNGANPLHEPYTGHHFPDGDMVHAYKFEGHSA-SYA 190	XX Disclosure: Fig 1; 135pp; English.
Db	132	ERPPVHELPSPGRIPFDPGYARANGANPFDPVGHLEFDGDMVHALTRNGAESTA 191	XX This invention relates to defence-related signalling genes isolated from
QY	191	CRFTOTNRFYQEROLGRPVPKAIGELHGHTGIAIRMLFLYARAAGTYDPAHGTGCVANG 250	XX the sunflower (<i>Helianthus annuus</i>). The genes encode a neoxanthin cleavage
Db	192	CRFTETARLQEERAIGRPVPKAIGELHGSGTIALYARAAGLYDPAHGTGCVANG 251	XX enzyme (NC), an amino acid permease (AAP) and a glutamic acid rich
QY	251	LYYFNGRLLAMSSEDDLPYQWQITPNGLKTVGREDFGDQLESTMIAHPKVDPERGELFAL 310	XX protein (GRP). The signalling gene is useful for increasing the
Db	252	LYYFNGRLLAMSSEDDLPYHVRVADGGDLTEVGRDFDQGLGCAMIAHPKLDPATGEHLAL 311	XX resistance of a plant to a pathogen such as fungus, virus, bacterium,
QY	311	SYDVSKPKTYFRSPDGTKSPDKVQDQVVFQVKLPMEI 370	XX nematode or insect (e.g. European corn borer), preferably
Db	312	SYDVKTPRQPTKYFVERPDGTSKSDDOVIEPLEQOPTMHDFAATENFVYVVDHQVVFQVKLQEMI 371	XX Sclerotinia spp., <i>Phoma</i> spp., or <i>Phomopsis</i> spp., by stably incorporating a
QY	371	RGGSPVVYDKNKVAFPGILDKYADEPSNIRWKIDAPDCFCFHLLNAWEPEPTDEYVWIGSC 430	XX construct containing the gene into the genome of the plant. The gene is
Db	372	RGGSPVVYDKREKTSRGFLKHAADASENAWVYDPCFCFHLLNAWEDEATGEGVVWIGSC 431	XX useful for regulating gene expression in a plant, in response to a
QY	431	MTPPDSEFNDSENLKSVLSEIRLNLKTTGESTTRPLISNEDQOQYNLEAGMVNRNNMLGRKT 490	XX stimulus such as infection with a pathogen, damage from a pathogen,
Db	432	MTPDSEFNDSERLESVLEIRLDPARTGESTRAVLP-FSQOENLEVQMVNRNLGREGS 490	XX hydrogen peroxide, jasmonic acid, methyl jasmonic acid, salicylic acid, oxalic acid or expression of a gene encoding oxalic acid oxidase. The
QY	491	KFAYLALAEWPWKYSFGEAKWVQVDTGEGVKHLYGDNRYGEGPLEPFGEGSE--- 545	XX genes are also useful for stem-preferred regulation of gene expression in
Db	491	RYAYLAVAEWPWKESGEAAEFLSTGELTKEYGRGEGPCFVPMDDPSAAHPRGEDSY 550	XX a plant. The genes are useful in agriculture, particularly in the
QY	546	ILCFVHDEKTWKSELQTYINAVSLVEATVLPSPRVYQIGTFIGADDILAKQ 597	XX breeding of crop plants with improved agronomic traits, for modifying amino acid (ABA) metabolism and for modulating amino acid transport and
Db	551	VLTVEVHDERAGTSELLVVAADIRLEATVLPSPRVFEGHTGFTITGQELAQ 602	XX content in plants. The present sequence represents a neoxanthin cleavage
QY	*	*	CC enzyme-like protein from <i>Zea</i> may be used in the characterisation of
Db	*	*	CC sunflower NC.
QY	*	*	XX
QY	604 AA;	Sequence	604 AA;
		Query Match	61.3%; Score 1930; DB 22; Length 604;
		Best Local Similarity	63.3%; Pred. No. 2.1e-186;
		Matches	375; Mismatches 73; Indels 22; Gaps 8;
QY	22	PPLSSQSSDLSYCSSLPMASRVTRKLNVSSALHTP---PALHFFPKQSSNSPATIVV--- 74	XX
Db	17	PARSARAS---NSVRSPRAYSSVPPCETIOFFPKYADLPAPSRRKPAAPAVGHA 71	XX
QY	75	-PKAKESNTQMNLEQRAAAAALDA-AEGFLVS-BERKLPPLPKTADPSVQIAGNEFAPVN 131	XX
Db	72	AAPRKAEGGKQLNLFQRAAAAALDAEFSFVANLERSPFGLPLSTADPAVQIAGNFAPVG 131	XX
QY	132	EQPVRNRLPVVGKLDPSIKGVYVRRNGANPLHEPYTGHHFPDGDMVHAYKFEGHSA-SYA 190	XX
Db	132	ERPPVHELPSPGRIPFDPGYARANGANPFDPVGHLEFDGDMVHALTRNGAESTA 191	XX
QY	191	CRFTOTNRFYQEROLGRPVPKAIGELHGHTGIAIRMLFLYARAAGTYDPAHGTGCVANG 250	XX
Db	192	CRFTETARLQEERAIGRPVPKAIGELHGSGTIALYARAAGLYDPAHGTGCVANG 251	XX
QY	251	LYYFNGRLLAMSSEDDLPYQWQITPNGLKTVGREDFGDQLESTMIAHPKVDPERGELFAL 310	XX
Db	252	LYYFNGRLLAMSSEDDLPYHVRVADGGDLTEVGRDFDQGLGCAMIAHPKLDPATGEHLAL 311	XX
QY	311	SYDVSKPKTYFRSPDGTKSPDKVQDQVVFQVKLPMEI 370	XX
Db	312	SYDVKTPRQPTKYFVERPDGTSKSDDOVIEPLEQOPTMHDFAATENFVYVVDHQVVFQVKLQEMI 371	XX
QY	371	RGGSPVVYDKNKVAFPGILDKYADEPSNIRWKIDAPDCFCFHLLNAWEPEPTDEYVWIGSC 430	XX
Db	372	RGGSPVVYDKREKTSRGFLKHAADASENAWVYDPCFCFHLLNAWEDEATGEGVVWIGSC 431	XX
QY	431	MTPPDSEFNDSENLKSVLSEIRLNLTGTGESTTRPLISNEDQOQYNLEAGMVNRNNMLGRKT 490	XX
Db	432	MTPDSEFNDSERLESVLEIRLDPARTGESTRAVLP-FSQOENLEVQMVNRNLGREGS 490	XX
QY	491	KFAYLALAEWPWKYSFGEAKWVQVDTGEGVKHLYGDNRYGEGPLEPFGEGSE--- 545	XX
Db	491	RYAYLAVAEWPWKESGEAAEFLSTGELTKEYGRGEGPCFVPMDDPSAAHPRGEDSY 550	XX
QY	546	ILCFVHDEKTWKSELQTYINAVSLVEATVLPSPRVYQIGTFIGADDILAKQ 597	XX
Db	551	VLTVEVHDERAGTSELLVVAADIRLEATVLPSPRVFEGHTGFTITGQELAQ 602	XX
QY	*	*	XX
Db	*	*	XX
QY	604 AA;	Sequence	604 AA;
		Query Match	61.3%; Score 1930; DB 22; Length 604;
		Best Local Similarity	63.3%; Pred. No. 2.1e-186;
		Matches	375; Mismatches 73; Indels 22; Gaps 8;
QY	22	PPLSSQSSDLSYCSSLPMASRVTRKLNVSSALHTP---PALHFFPKQSSNSPATIVV--- 74	XX
Db	17	PARSARAS---NSVRSPRAYSSVPPCETIOFFPKYADLPAPSRRKPAAPAVGHA 71	XX
QY	75	-PKAKESNTQMNLEQRAAAAALDA-AEGFLVS-BERKLPPLPKTADPSVQIAGNEFAPVN 131	XX
Db	72	AAPRKAEGGKQLNLFQRAAAAALDAEFSFVANLERSPFGLPLSTADPAVQIAGNFAPVG 131	XX
QY	132	EQPVRNRLPVVGKLDPSIKGVYVRRNGANPLHEPYTGHHFPDGDMVHAYKFEGHSA-SYA 190	XX
Db	132	ERPPVHELPSPGRIPFDPGYARANGANPFDPVGHLEFDGDMVHALTRNGAESTA 191	XX
QY	191	CRFTOTNRFYQEROLGRPVPKAIGELHGHTGIAIRMLFLYARAAGTYDPAHGTGCVANG 250	XX
Db	192	CRFTETARLQEERAIGRPVPKAIGELHGSGTIALYARAAGLYDPAHGTGCVANG 251	XX
QY	251	LYYFNGRLLAMSSEDDLPYQWQITPNGLKTVGREDFGDQLESTMIAHPKVDPERGELFAL 310	XX
Db	252	LYYFNGRLLAMSSEDDLPYHVRVADGGDLTEVGRDFDQGLGCAMIAHPKLDPATGEHLAL 311	XX
QY	311	SYDVSKPKTYFRSPDGTKSPDKVQDQVVFQVKLPMEI 370	XX
Db	312	SYDVKTPRQPTKYFVERPDGTSKSDDOVIEPLEQOPTMHDFAATENFVYVVDHQVVFQVKLQEMI 371	XX
QY	371	RGGSPVVYDKNKVAFPGILDKYADEPSNIRWKIDAPDCFCFHLLNAWEPEPTDEYVWIGSC 430	XX
Db	372	RGGSPVVYDKREKTSRGFLKHAADASENAWVYDPCFCFHLLNAWEDEATGEGVVWIGSC 431	XX
QY	431	MTPPDSEFNDSENLKSVLSEIRLNLTGTGESTTRPLISNEDQOQYNLEAGMVNRNNMLGRKT 490	XX
Db	432	MTPDSEFNDSERLESVLEIRLDPARTGESTRAVLP-FSQOENLEVQMVNRNLGREGS 490	XX
QY	491	KFAYLALAEWPWKYSFGEAKWVQVDTGEGVKHLYGDNRYGEGPLEPFGEGSE--- 545	XX
Db	491	RYAYLAVAEWPWKESGEAAEFLSTGELTKEYGRGEGPCFVPMDDPSAAHPRGEDSY 550	XX
QY	546	ILCFVHDEKTWKSELQTYINAVSLVEATVLPSPRVYQIGTFIGADDILAKQ 597	XX
Db	551	VLTVEVHDERAGTSELLVVAADIRLEATVLPSPRVFEGHTGFTITGQELAQ 602	XX
QY	*	*	XX
Db	*	*	XX
QY	604 AA;	Sequence	604 AA;
		Query Match	61.3%; Score 1930; DB 22; Length 604;
		Best Local Similarity	63.3%; Pred. No. 2.1e-186;
		Matches	375; Mismatches 73; Indels 22; Gaps 8;
QY	22	PPLSSQSSDLSYCSSLPMASRVTRKLNVSSALHTP---PALHFFPKQSSNSPATIVV--- 74	XX
Db	17	PARSARAS---NSVRSPRAYSSVPPCETIOFFPKYADLPAPSRRKPAAPAVGHA 71	XX
QY	75	-PKAKESNTQMNLEQRAAAAALDA-AEGFLVS-BERKLPPLPKTADPSVQIAGNEFAPVN 131	XX
Db	72	AAPRKAEGGKQLNLFQRAAAAALDAEFSFVANLERSPFGLPLSTADPAVQIAGNFAPVG 131	XX
QY	132	EQPVRNRLPVVGKLDPSIKGVYVRRNGANPLHEPYTGHHFPDGDMVHAYKFEGHSA-SYA 190	XX
Db	132	ERPPVHELPSPGRIPFDPGYARANGANPFDPVGHLEFDGDMVHALTRNGAESTA 191	XX
QY	191	CRFTOTNRFYQEROLGRPVPKAIGELHGHTGIAIRMLFLYARAAGTYDPAHGTGCVANG 250	XX
Db	192	CRFTETARLQEERAIGRPVPKAIGELHGSGTIALYARAAGLYDPAHGTGCVANG 251	XX
QY	251	LYYFNGRLLAMSSEDDLPYQWQITPNGLKTVGREDFGDQLESTMIAHPKVDPERGELFAL 310	XX
Db	252	LYYFNGRLLAMSSEDDLPYHVRVADGGDLTEVGRDFDQGLGCAMIAHPKLDPATGEHLAL 311	XX
QY	311	SYDVSKPKTYFRSPDGTKSPDKVQDQVVFQVKLPMEI 370	XX
Db	312	SYDVKTPRQPTKYFVERPDGTSKSDDOVIEPLEQOPTMHDFAATENFVYVVDHQVVFQVKLQEMI 371	XX
QY	371	RGGSPVVYDKNKVAFPGILDKYADEPSNIRWKIDAPDCFCFHLLNAWEPEPTDEYVWIGSC 430	XX
Db	372	RGGSPVVYDKREKTSRGFLKHAADASENAWVYDPCFCFHLLNAWEDEATGEGVVWIGSC 431	XX
QY	431	MTPPDSEFNDSENLKSVLSEIRLNLTGTGESTTRPLISNEDQOQYNLEAGMVNRNNMLGRKT 490	XX
Db	432	MTPDSEFNDSERLESVLEIRLDPARTGESTRAVLP-FSQOENLEVQMVNRNLGREGS 490	XX
QY	491	KFAYLALAEWPWKYSFGEAKWVQVDTGEGVKHLYGDNRYGEGPLEPFGEGSE--- 545	XX
Db	491	RYAYLAVAEWPWKESGEAAEFLSTGELTKEYGRGEGPCFVPMDDPSAAHPRGEDSY 550	XX
QY	546	ILCFVHDEKTWKSELQTYINAVSLVEATVLPSPRVYQIGTFIGADDILAKQ 597	XX
Db	551	VLTVEVHDERAGTSELLVVAADIRLEATVLPSPRVFEGHTGFTITGQELAQ 602	XX
QY	*	*	XX
Db	*	*	XX
QY	604 AA;	Sequence	604 AA;
		Query Match	61.3%; Score 1930; DB 22; Length 604;
		Best Local Similarity	63.3%; Pred. No. 2.1e-186;
		Matches	375; Mismatches 73; Indels 22; Gaps 8;
QY	22	PPLSSQSSDLSYCSSLPMASRVTRKLNVSSALHTP---PALHFFPKQSSNSPATIVV--- 74	XX
Db	17	PARSARAS---NSVRSPRAYSSVPPCETIOFFPKYADLPAPSRRKPAAPAVGHA 71	XX
QY	75	-PKAKESNTQMNLEQRAAAAALDA-AEGFLVS-BERKLPPLPKTADPSVQIAGNEFAPVN 131	XX
Db	72	AAPRKAEGGKQLNLFQRAAAAALDAEFSFVANLERSPFGLPLSTADPAVQIAGNFAPVG 131	XX
QY	132	EQPVRNRLPVVGKLDPSIKGVYVRRNGANPLHEPYTGHHFPDGDMVHAYKFEGHSA-SYA 190	XX
Db	132	ERPPVHELPSPGRIPFDPGYARANGANPFDPVGHLEFDGDMVHALTRNGAESTA 191	XX
QY	191	CRFTOTNRFYQEROLGRPVPKAIGELHGHTGIAIRMLFLYARAAGTYDPAHGTGCVANG 250	XX
Db	192	CRFTETARLQEERAIGRPVPKAIGELHGSGTIALYARAAGLYDPAHGTGCVANG 251	XX
QY	251	LYYFNGRLLAMSSEDDLPYQWQITPNGLKTVGREDFGDQLESTMIAHPKVDPERGELFAL 310	XX
Db	252	LYYFNGRLLAMSSEDDLPYHVRVADGGDLTEVGRDFDQGLGCAMIAHPKLDPATGEHLAL 311	XX
QY	311	SYDVSKPKTYFRSPDGTKSPDKVQDQVVFQVKLPMEI 370	XX
Db	312	SYDVKTPRQPTKYFVERPDGTSKSDDOVIEPLEQOPTMHDFAATENFVYVVDHQVVFQVKLQEMI 371	XX
QY	371	RGGSPVVYDKNKVAFPGILDKYADEPSNIRWKIDAPDCFCFHLLNAWEPEPTDEYVWIGSC 430	XX
Db	372	RGGSPVVYDKREKTSRGFLKHAADASENAWVYDPCFCFHLLNAWEDEATGEGVVWIGSC 431	XX
QY	431	MTPPDSEFNDSENLKSVLSEIRLNLTGTGESTTRPLISNEDQOQYNLEAGMVNRNNMLGRKT 490	XX
Db	432	MTPDSEFNDSERLESVLEIRLDPARTGESTRAVLP-FSQOENLEVQMVNRNLGREGS 490	XX
QY	491	KFAYLALAEWPWKYSFGEAKWVQVDTGEGVKHLYGDNRYGEGPLEPFGEGSE--- 545	XX
Db	491	RYAYLAVAEWPWKESGEAAEFLSTGELTKEYGRGEGPCFVPMDDPSAAHPRGEDSY 550	XX
QY	546	ILCFVHDEKTWKSELQTYINAVSLVEATVLPSPRVYQIGTFIGADDILAKQ 597	XX
Db	551	VLTVEVHDERAGTSELLVVAADIRLEATVLPSPRVFEGHTGFTITGQELAQ 602	XX
QY	*	*	XX
Db	*	*	XX
QY	604 AA;	Sequence	604 AA;
		Query Match	61.3%; Score 1930; DB 22; Length 604;
		Best Local Similarity	63.3%; Pred. No. 2.1e-186;
		Matches	375; Mismatches 73; Indels 22; Gaps 8;
QY	22	PPLSSQSSDLSYCSSLPMASRVTRKLNVSSALHTP---PALHFFPKQSSNSPATIVV--- 74	XX
Db	17	PARSARAS---NSVRSPRAYSSVPPCETIOFFPKYADLPAPSRRKPAAPAVGHA 71	XX
QY	75	-PKAKESNTQMNLEQRAAAAALDA-AEGFLVS-BERKLPPLPKTADPSVQIAGNEFAPVN 131	XX
Db	72	AAPRKAEGGKQLNLFQRAAAAALDAEFSFVANLERSPFGLPLSTADPAVQIAGNFAPVG 131	XX
QY	132	EQPVRNRLPVVGKLDPSIKGVYVRRNGANPLHEPYTGHHFPDGDMVHAYKFEGHSA-SYA 190	XX
Db	132	ERPPVHELPSPGRIPFDPGYARANGANPFDPVGHLEFDGDMVHALTRNGAESTA 191	XX
QY	191	CRFTOTNRFYQEROLGRPVPKAIGELHGHTGIAIRMLFLYARAAGTYDPAHGTGCVANG 250	XX
Db	192	CRFTETARLQEERAIGRPVPKAIGELHGSGTIALYARAAGLYDPAHGTGCVANG 251	XX
QY	251	LYYFNGRLLAMSSEDDLPYQWQITPNGLKTVGREDFGDQLESTMIAHPKVDPERGELFAL 310	XX
Db	252	LYYFNGRLLAMSSEDDLPYHVRVADGGDLTEVGRDFDQGLGCAMIAHPKLDPATGEHLAL 311	XX
QY	311	SYDVSKPKTYFRSPDGTKSPDKVQDQVVFQVKLPMEI 370	XX
Db	312	SYDVKTPRQPTKYFVERPDGTSKSDDOVIEPLEQOPTMHDFAATENFVYVVDHQVVFQVKLQEMI 371	XX
QY	371	RGGSPVVYDKNKVAFPGILDKYADEPSNIRWKIDAPDCFCFHLLNAWEPEPTDEYVWIGSC 430	XX
Db	372	RGGSPVVYDKREKTSRGFLKHAADASENAWVYDPCFCFHLLNAWEDEATGEGVVWIGSC 431	XX
QY	431	MTPPDSEFNDSENLKSVLSEIRLNLTGTGESTTRPLISNEDQOQYNLEAGMVNRNNMLGRKT 490	XX
Db	432	MTPDSEFNDSERLESVLEIRLDPARTGESTRAVLP-FSQOENLEVQMVNRNLGREGS 490	XX
QY	491	KFAYLALAEWPWKYSFGEAKWVQVDTGEGVKHLYGDNRYGEGPLEPFGEGSE--- 545	XX
Db	491	RYAYLAVAEWPWKESGEAAEFLSTGELTKEYGRGEGPCFVPMDDPSAAHPRGEDSY 550	XX
QY	546	ILCFVHDEKTWKSELQTYINAVSLVEATVLPSPRVYQIGTFIGADDILAKQ 597	XX
Db	551	VLTVEVHDERAGTSELLVVAADIRLEATVLPSPRVFEGHTGFTITGQELAQ 602	XX
QY	*	*	XX
Db	*	*	XX
QY	604 AA;	Sequence	604 AA;
		Query Match	61.3%; Score 1930; DB 22; Length 604;
		Best Local Similarity	63.3%; Pred. No. 2.1e-186;
		Matches	375; Mismatches 73; Indels 22; Gaps 8;
QY	22	PPLSSQSSDLSYCSSLPMASRVTRKLNVSSALHTP---PALHFFPKQSSNSPATIVV--- 74	XX
Db	17	PARSARAS---NSVRSPRAYSSVPPCETIOFFPKYADLPAPSRRKPAAPAVGHA 71	XX
QY	75	-PKAKESNTQMNLEQRAAAAALDA-AEGFLVS-BERKLPPLPKTADPSVQIAGNEFAPVN 131	XX
Db	72	AAPRKAEGGKQLNLFQRAAAAALDAEFSFVANLERSPFGLPLSTADPAVQIAGNFAPVG 131	XX
QY	132	EQPVRNRLPVVGKLDPSIKGVYVRRNGANPLHEPYTGHHFPDGDMVHAYKFEGHSA-SYA 190	XX
Db	132	ERPPVHELPSPGRIPFDPGYARANGANPFDPVGHLEFDGDMVHALTRNGAESTA 191	XX
QY	191	CRFTOTNRFYQEROLGRPVPKAIGELHGHTGIAIRMLFLYARAAGTYDPAHGTGCVANG 250	XX
Db	192	CRFTETARLQEERAIGRPVPKAIGELHGSGTIALYARAAGLYDPAHGTGCVANG 251	XX
QY	251	LYYFNGRLLAMSSEDDLPYQWQITPNGLKTVGREDFGDQLESTMIAHPKVDPERGELFAL 310	XX
Db	252	LYYFNGRLLAMSSEDDLPYHVRVADGGDLTEVGRDFDQGLGCAMIAHPKLDPATGEHLAL 311	XX
QY	311	SYDVSKPKTYFRSPDGTKSPDKVQDQVVFQVKLPMEI 370	XX
Db	312	SYDVKTPRQPTKYFVERPDGTSKSDDOVIEPLEQOPTMHDFAATENFVYVVDHQVVFQVKLQEMI 371	XX
QY	371	RGGSPVVYUUNKKVAFQVDFLKYADPSNKKWIDAPGDFCOPFLNNAWEPEPETDEVVVGSC 430	XX
Db	372	RGGSPVLLKEKTSRGFLPKBDAASENAWVDPDGFPSAGVVAAG 431	XX
QY	431	MTPPDSEFNDSENLKSVLSEIRLNLTGTGESTTRPLISNEDQOQYNLEAGMVNRNNMLGRKT 490	XX
Db	432	MTPDSEFNDSERLESVLEIRLDPARTGESTRAVLP-PSQQENLYGVMNNLNGRES 490	XX
QY	491	KFAYLALAEWPWKYSFGEAKWVQVDTGEGVKHLYGDNRYGEGPLEPFLGEGGE--- 545	XX
Db	491	RYAYLAVAEWPWKESGEAAEFLSTGELTKEYGRGEGPCFVPMDDPSAAHPRGEDSY 550	XX
QY	546	ILCFVHDEKTWKSELQTYINAVSLVEATVLPSPRVYQIGTFIGADDILAKQ 597	XX
Db	551	VLTVEVHDERAGTSELLVVAADIRLEATVLPSPRVFEGHTGFTITGQELAQ 602	XX
QY	*	*	XX
Db	*	*	XX
QY	604 AA;	Sequence	6

Db	491 RAYLAYAEPWPKESGFAKEIDLSTGELTKPEYGEGRFGEPCVPMDDAAAHPRGEDDGY	550	Qy	83 TKQMNLFQRAAAALDAAE-GFLYSHEKLHPLKTTADPSVQAGNFPAPVNEQPVVRNLPV	141
Qy	546 ILCEVHDERTWKSELQITQINAVASLVEATVKLPSRVPYGEHGTPIGADDLAKQ	597	Db	63 ---INLLQKLAAMLDKKESSIVIPMEONRPLKPTDAAVLSGNFAVNECPVQNGLEV	119
Db	551 VLTFFVHDERAGTISSELLVNAADIRLEATVQLPSRVPFGFHGGTFITGQLEAQ	602	Qy	142 VKGLPDSIKGVYVTRNGANPLHEPYTGHHFFDGDMVHAA--FEHGSASYACRFQTQINRF	199
RESULT 10			Db	120 VGO1PSCKKGVYTRNGANPMPEPLAGHILFDGSMIAHVASIGFDN-QWSYSCRYTKNRL	178
AAB04786	AAE04786 standard; Protein; 577 AA.		Qy	200 VQERLGPVFPVPAIGEJLGHITGTLRMLVPAHGTVYANAGLVYFNERRL	259
XX			Db	179 VQETALGVSVPKPGIHLGHQSCLARALFTARAGIVLVDTRGMGYANAGVVFNRRL	238
AC			Qy	260 AMSEDDLFVQVQQTIPGDLKTYGRDFDQCLESTMIAHPKVDPESGEFALSYDVSKPV	319
XX			Db	239 LKYFRFSPDGTKSPDVEQDQOPTMMHDEATPENFVYVPDQQVVFKLPEMIRGGSPVYD	435
DT			Qy	320 LRYLKFNCGKTRDVEITLPEPTMHEFAITENFVYVLDQOMVFKLSEMIRGGSPVY	379
DE			Db	259 LRYLKFNCGKTRDVEITLPEPTMHEFAITENFVYVLDQOMVFKLSEMIRGGSPVY	358
Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED5			Qy	380 KRYVARFGLDKYAEDSSNIKHDAPDCEFLWNAAVEPEDE---VVVVGSCMPD	473
XX			Db	359 KEMMARFVLSQDLTSDDINWVDPFCFCFLWNAAVE-RTEEGDFVTVVVGSCMPD	417
Neoxanthin cleavage enzyme; AtNCED5; abscisic acid; ABA; herbicide;			Qy	436 SIFNESDENLKSVLSETRINKTGESTRPPIISNEDQVNLEAGMNRNLGRKTIFAYL	495
KW stress tolerance; transgenic plant; Plant breeding; antisense therapy;			Db	418 TIFSESEGPTRVYELRNLMTIKESNRKVLYTG---VNLFAUGHNSYGR&SOFVYI	473
KW plant growth protectant.			Qy	496 ALAEPWPKVSGPAKVDLTTGEVKHLGYDNRYGGEPFLPPEGGEDEGGYLLCFVHDEKT	555
XX			Db	474 AIAADPKCSCGTAKVDFQNTGVSEFNYGSPSRGGEPQPVPSEGEEDKGYVMGFVREEK	533
OS Arabidopsis thaliana.			Qy	556 WKSELQIVNAVSELEATVKLPSRVPYGFHGTFIGADDLAKQ	598
XX			Db	534 DESEFVYVDAIDTMKQVAANRPERVPYGFHGTIVSENQNLKEQV	576
EP1116794-A2.			RESULT 11		
XX			ABB92416		
PD 18-JUL-2001.			ID		
XX			ABB92416 standard; Protein; 577 AA.		
PD 11-JAN-2001; 2001IEP-0300218.			XX		
XX			AC		
PR 13-JAN-2000; 2000JUP-0010056.			XX		
PR 11-JAN-2001; 2001JUP-0003476.			XX		
XX			XX		
PA (RIKE) RIKEN KK.			XX		
XX			XX		
P1 Iuchi S, Kobayashi M, Shinozaki K;			XX		
XX			XX		
WPI; 2001-400081/43.			XX		
DR N-PSDB; AAD09398			XX		
XX			XX		
PT A DNA encoding a protein with a neoxanthin cleavage activity for			XX		
PT producing transgenic plants with improved or decreased stress tolerance			XX		
PT XX			XX		
PS Claim 3; Fig 10; 101PP; English.			XX		
PS XX			XX		
PT The invention relates to neoxanthin cleavage enzymes and their			XX		
PT corresponding cDNA molecules. Neoxanthin cleavage enzymes play a key			XX		
PT role in endogenous abscisic acid (ABA) biosynthesis under drought stress.			XX		
PT Neoxanthin cleavage enzyme is used for improving stress tolerance in a			XX		
PT plant when expressed in a plant cell. The invention also relates to			XX		
PT methods for increasing or decreasing stress tolerance in a plant by			XX		
PT introducing the DNA into the plant, and a transgenic plant into which a			XX		
PT neoxanthin cleavage enzyme is introduced. The improvement of stress			XX		
PT tolerance in plants is useful, for example in plant breeding. Neoxanthin			XX		
PT cleavage enzyme genes are useful for producing transgenic plants. An arid			XX		
PT land can be improved by growing transgenic weed for several years and			XX		
PT then removing the weed by specifically lowering stress tolerance in the			XX		
PT weed by inducing an inducible promoter. The present sequence is			XX		
PT Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED5, protein.			XX		
CC AtNCED5 cDNA is obtained from an Arabidopsis plant-derived cDNA			XX		
CC library using a cDNA of the CPRD65 (Cowpea Responsive to Dehydration)			XX		
CC gene isolated from cowpea plant as a probe.			XX		
SQ Sequence 577 AA;			XX		
Query Match Score 1663.5; DB 22; Length 577;			XX		
Best Local Similarity 55.4%; Pred. No. 2.2e-159;			XX		
Matches 323; Conservative 100; Mismatches 137; Indels 23; Gaps 9;			XX		
Qy 24 LSLOSSDLSYCSLSEPMASRVKLNSS-ALHHTPALHFPKQSSNSPAVVKPAKESN 82			XX		
Db 9 LIPTKTKTSPRSHLLPQPKNANIRRILINPFKIPILPDLISPVPS-----VKLKPTYPN-62			XX		
The invention relates to identifying target proteins			CC		

(ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides.

XX Sequence 577 AA;

Query Match 52.8%; Score 1663.5; DB 23; Length 577;

Best Local Similarity 55.4%; Pred. No. 2.2e-159; Mismatches 137; Index 23; Gaps 9;

Qy 24 LSSQSSDLSCSSLPMASRYTRKINVSS-ALHTPPALHPKQSSNSPATVVKPRAKESN 82

Db 9 LLPTKISPRSHLFLPKRNNSRRLINPFKPTIPLDLTSVPSP ---- -VKLKPTYPN 62

Qy 83 TKOMNLFORAAAALDAAE-GFLYSHKEKLHLPLKPTADPSQIAGNFAPNEQPVFRNLPV 141

Db 63 -- -LNLLQKIAATMLDKIESSTIVIPMEQNRLPLKPTDPAVLGCFNAPVNECPVONGLEV 119

Qy 142 VGPLPDSTIKGYVVRNGANP-LHEPVTHFVHFEDGMVHAYRK -- FEHGSASYACRFQTNRF 199

Db 120 VQQLISCLSKVYIIRGANPMEPLAHHLDGDMIHAVSIGFDN_QVSYSCTRYTKTNRL 178

Qy 200 VQEROLGRPFPPKAIGELHHTGTLARLMFYARAAGTVPAHNGVANGNGLYFNGRLI 259

Db 179 VQETALGRSYSEPKPAGELHHSGLARLARASIGLVGTRMVGVANGVVFNGRL 238

Qy 260 AMSEDDLPVQYQITENGDLKTVGRDFDGSLESTMIAHKPVDPESGEELFALSYDVWSKY 319

Db 239 AMSEDDLPVYKIDQCGDLETTIGRFQFDQDIDSVTAHKVDAATGDLHILSYNVLKPH 298

Qy 320 LKYFRESPDCTKSPDVEIQDQPMHMHDFAITENFVVVDQOYVVEKLPMEMIRGGSPVYYD 379

Db 299 LRYLKFNTGKKTRDVEITPEPIMHDFAITENFVVIDQQMYFKLSEMIRGGSPVIV 358

Qy 380 KNYTAFRGFLDKYAKEDSSNNKWDIAPDCIFQFLNAAWEPEPTDE --- VVVGSCMTPD 435

Db 359 KEMMARFGVLSKQDLTGSDINWVDPDCECFHLNAAWEETTEGDPVIVVIGCMSPD 417

Qy 436 STINESDENIKSVLSEIRIUNLKGETSTRREPILSNDQOYNNLEAGMVNRMLGRKTKEYL 495

Db 418 TIEFEGEPEPTRLSEIRIUNRTEKESNRKVTVTG --- VNLEAHHINRYVGRSQFYI 473

Qy 496 ALAPWPWKYSGEAKVQLTGEVKKHYGDNRGCGEPLFPLGEGSEEDEGYVLCFVHDBKT 555

Db 474 AIADWPWKCSGIAKVQDINGTVSEFNFGSREFPCFVPEGSEEDKGYVMGFSVREDEK 533

Qy 556 WKSLOQIYNAVSLFEVATVYKLPSPVYQHGHTFVGADDIJKQV 598

Db 534 DESEFVVVDTDKQVAAVRLPERPVYQGFFGTFSVSENQIKEV 576

RESULT 12

AAG11333 standard; Protein; 595 AA.

ID AAG11333

XX AC AAG11333;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 37612.

XX KW Protein identification; Signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX	06-SEP-2000.	PD	25-FEB-2000; 2000EP-0301439.
XX		PP	25-MAR-1999; 990S-0121825.
XX		PR	05-MAR-1999; 990S-0123180.
XX		PR	09-MAR-1999; 990S-0123548.
XX		PR	23-MAR-1999; 990S-0125788.
XX		PR	25-MAR-1999; 990S-0126264.
XX		PR	29-MAR-1999; 990S-0126785.
XX		PR	01-APR-1999; 990S-0127462.
XX		PR	06-APR-1999; 990S-0128234.
XX		PR	08-APR-1999; 990S-0128714.
XX		PR	16-APR-1999; 990S-0129845.
XX		PR	19-APR-1999; 990S-0130077.
XX		PR	21-APR-1999; 990S-0130449.
XX		PR	23-APR-1999; 990S-0130510.
XX		PR	23-APR-1999; 990S-0130891.
XX		PR	16-APR-1999; 990S-01311449.
XX		PR	18-APR-1999; 990S-0132048.
XX		PR	20-APR-1999; 990S-0132407.
XX		PR	04-MAY-1999; 990S-0132484.
XX		PR	05-MAY-1999; 990S-0132485.
XX		PR	06-MAY-1999; 990S-0132486.
XX		PR	07-MAY-1999; 990S-0132487.
XX		PR	11-MAY-1999; 990S-0132833.
XX		PR	14-MAY-1999; 990S-0132426.
XX		PR	14-MAY-1999; 990S-0132428.
XX		PR	14-MAY-1999; 990S-0134218.
XX		PR	14-MAY-1999; 990S-0134219.
XX		PR	14-MAY-1999; 990S-0134221.
XX		PR	14-MAY-1999; 990S-0134370.
XX		PR	18-MAY-1999; 990S-0134768.
XX		PR	19-MAY-1999; 990S-0134940.
XX		PR	20-MAY-1999; 990S-0135724.
XX		PR	21-MAY-1999; 990S-0135725.
XX		PR	24-MAY-1999; 990S-0135633.
XX		PR	25-MAY-1999; 990S-0136021.
XX		PR	27-MAY-1999; 990S-0136392.
XX		PR	28-MAY-1999; 990S-0136782.
XX		PR	01-JUN-1999; 990S-0137222.
XX		PR	03-JUN-1999; 990S-0137228.
XX		PR	04-JUN-1999; 990S-0137502.
XX		PR	07-JUN-1999; 990S-0137724.
XX		PR	08-JUN-1999; 990S-0138994.
XX		PR	10-JUN-1999; 990S-0138440.
XX		PR	14-JUN-1999; 990S-0138847.
XX		PR	16-JUN-1999; 990S-0139119.
XX		PR	16-JUN-1999; 990S-0139452.
XX		PR	17-JUN-1999; 990S-0139453.
XX		PR	17-JUN-1999; 990S-0139454.
XX		PR	18-JUN-1999; 990S-0139455.
XX		PR	18-JUN-1999; 990S-0139457.
XX		PR	18-JUN-1999; 990S-0139458.
XX		PR	18-JUN-1999; 990S-0139460.
XX		PR	18-JUN-1999; 990S-0139461.
XX		PR	18-JUN-1999; 990S-0139462.
XX		PR	18-JUN-1999; 990S-0139463.
XX		PR	18-JUN-1999; 990S-0139465.
XX		PR	18-JUN-1999; 990S-0139467.
XX		PR	21-JUN-1999; 990S-0139468.
XX		PR	22-JUN-1999; 990S-0139469.
XX		PR	23-JUN-1999; 990S-0139470.
XX		PR	24-JUN-1999; 990S-0139471.
XX		PR	28-JUN-1999; 990S-0140823.
XX		PR	29-JUN-1999; 990S-0140991.
XX		PR	30-JUN-1999; 990S-0141287.
XX		PR	01-JUL-1999; 990S-0141842.
XX		PR	01-JUL-1999; 990S-0142154.

BEST AVAILABLE COPY

PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144055.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144335.
 PR 22-JUL-1999; 99US-0144532.
 PR 22-JUL-1999; 99US-0144532.
 PR 22-JUL-1999; 99US-0145089.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-01445085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145221.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146385.
 PR 02-AUG-1999; 99US-0146386.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147200.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148644.
 PR 16-AUG-1999; 99US-0149002.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149416.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149773.
 PR 20-AUG-1999; 99US-0149939.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150366.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151530.
 PR 07-SEP-1999; 99US-0152265.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153158.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0156559.
 PR 28-SEP-1999; 99US-016458.
 PR 29-SEP-1999; 99US-016596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 18-OCT-1999; 99US-0159295.
 PR 18-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 13-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 22-OCT-1999; 99US-0160815.
 PR 21-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 30.7%; Score 968; DB 21; Length 595;
 Best Local Similarity 36.7%; Pred. No. 1e-88;
 Matches 210; Conservative 106; Mismatches 220; Indels 36; Gaps 13;

QY 4.2 SRVIRKLNVSSALHTPPALHPKQSSNSPAIVVKPKAKESNTKOMNLFLORAAAAAALDAAE 101
 Db 4.2 SPITNPSDNNDRRNPKTBLH --NRTNTLVLSSPPKLRPLMTATLF ---TTVEDVIN 94
 QY 10.2 GFLVSHEKLHPLPKTADPSVQIAGNAPVNEQPVVRNLNPVY-GKLPSDIIKGVVYRNGANP 160
 Db 95 T-FIDP --- -PSRPSPDVKHLSDNPAVPLDDELPPTDCETIHGTPLSLNGAYTINGPNP 149
 QY 16.1 LHEPVTHGHFFGDGMVHYAKFENGSSASYACRFQTINRPFYERQFLGRPFYFPKAIGELHGH 220
 Db 150 QFLPRGPYHIFDGDGMHLAIKHNSKATLCSRVTYKTVNEKQTGAPMPNVTSFGFNGV 209
 QY 22.1 T-GIARLMFYARAAGIVDPAHGTGVANAGLVYENGRLLAMSEDDLPYQVOITTPNGDLK 279
 Db 210 TASVARGATAARVLTGQYNPVANGLANSFLAASSNRFAFLGSDLPPAVRLTESGDE 269
 QY 28.0 TVGRFDFDGOLESTMAHAKVDPESGEFALSVDYSPKSYLYKFRESPDTKSBDVEI Q 338
 Db 270 TIGRDFDGGGLAMSNTAHPTDPTGETAFRYGP -PFPLTFEREDSGAKKQRDVP/F/S 328
 QY 33.9 LDQPTMMHDEAITENMVFVYDQQVFK - -LPEMIRGGSPVYDQKVNKFQGFLDKYAFD 395
 Db 329 MTSPSFLHDEFAITKRHAIFAEIQGMRMNMFLVIEGSPVGIDNGKTFRLGTVPKYAGD 388
 QY 39.6 SSNIKWIADPCFCFHLMWAAEPTDEVVIGSCMTPPDSIFNESDENLKSVLSETRN 455
 Db 38.9 ESEMKWEEPGFNITHAINAWDEDGNSVTLAPNIMSSLEHTLERMD-LVHALVEKVKID 447
 QY 45.6 LKTGSTIRRPLISNEQQYNLEAGMVNRMLGRKTKFAYLALAEPWPVSKGAKVDTTG 515

OS	Arabidopsis thaliana.
XX	
PN	EP1116794-A2.
XX	
PD	18-JUL-2001.
XX	
PF	11-JAN-2001; 2001EP-0300218.
XX	
PR	13-JAN-2000; 2000JP-0010056.
PR	11-JAN-2001; 2001JP-0003476.
XX	
PA	(RIKE) RIKEN KK.
XX	
PI	Iuchi S, Kobayashi M, Shinozaki K;
XX	
DR	WPL: 2001-40008143.
DR	N-PADB: AAD0395.
XX	
PT	A DNA encoding a protein with a neoxanthin cleavage activity for precursors.

The invention relates to neoxanthin cleavage enzymes and their corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key role in endogenous abscisic acid (ABA) biosynthesis under drought stress. Neoxanthin cleavage enzyme is used for improving stress tolerance in a plant when expressed in a plant cell. The invention also relates to methods for increasing or decreasing stress tolerance in a plant by introducing the DNA into the plant, and a transgenic plant into which a neoxanthin cleavage enzyme is introduced. The improvement of stress tolerance in plants is useful, for example in plant breeding. Neoxanthin cleavage enzyme genes are useful for producing transgenic plants. An arable land can be improved by growing transformant weed for several years and then removing the weed by specifically lowering stress tolerance in the weed by inducing an inducible promoter. The present sequence is Arabidopsis thaliana neoxanthin cleavage enzyme, ATNCED2 protein. The ATNCED2 cDNA is obtained from an *Arabidopsis* plant-derived cDNA library using a cDNA from the CPT655 (Cowpea Responsive to Dehydration gene isolated from cowpea cDNA library).

XX Novel isolated defence-related signalling gene isolated from sunflower
 PT encoding neoxanthin cleavage enzyme, amino acid permease or glutamic
 acid rich protein useful for increasing resistance of plant to a
 pathogen
 XX Example 1; Fig 1; 135pp; English.
 XX

This invention relates to defence-related signalling genes isolated from the sunflower (Helianthus annuus). The genes encode a neoxanthin cleavage enzyme (NCE), an amino acid permease (AAP) and a glutamic acid rich protein (GRP). The signalling gene is useful for increasing the resistance of a plant to a pathogen such as fungus, virus, bacterium, nematode or insect (e.g. European corn borer), preferentially Schizentria spp., or Phomopsis spp., by stably incorporating a construct containing the gene into the genome of the plant. The gene is useful for regulating gene expression in a plant, in response to a stimulus such as infection with a pathogen, damage from a pathogen, hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid, oxalic acid or expression of a gene encoding oxalate oxidase. The genes are also useful for stem-preferred regulation of gene expression in a plant. The genes are useful in agriculture, particularly in the breeding of crop plants with improved agronomic traits, for modifying abscisic acid (ABA) metabolism and for modifying amino acid transport and content in plants. The present sequence represents a neoxanthin cleavage enzyme-like protein from *Arabidopsis thaliana*, used in the characterisation of sunflower NCE.

Sequence 595 AA;

Query Match 30.7%; Score 968; DB 22; Length 595;

Best Local Similarity 36.7%; Pred. No. 1e-88; Mismatches 220; Indels 36; Gaps 13;

Matches 210; Conservative 106; No. 1e-88;

Db 4.2 SFYTRKLKVSSALHTEPPALHPKQKSSNSPAVVKPRAKESNTKQMNLEQRAAAALDAAE 101

Db 4.2 SITNP-DNNRNRKPTLH--NRFNHTLYSSPPKRPEMTLATLF---TTVEDVIN 94

QY 102 GFUVSHEKHLHPLKPTADPSVQTLAGNFAPVNQPVQRVRNLPPV 160

Db 95 TFDIP-----PSRSPVDPKHVSNDNAPVLDLPEPDCEIITHGTPLSLNGAYIRGPNP 149

QY 161 LIEPTVGHFFEDGDMHVAYKEFHGSASYACRFQTQNRFVQEORQLGRPVFPKAIGELGH 220

Db 150 QFLPRSPYHLEFDGMLHAIKTHNGKATLCSSRYVKTQYKNEVKQFQAPYMFNVFGNGV 209

Qy 221 T-GIAHMLFVRAAAGIVDEPAHGTCVANGLVYNGRLAMSED DLLPYQVQITPNGDK 279

Db 210 TASVARGALTAARVLTGQYNNVNGIGLANTISLAFAGSNDPLPAVRLTSESGDIE 269

QY 280 TYGRERFDGQESTM TAHPKVDPESGELEFALS YDVSKPKYLKFRTSPDTKSPDYEI-Q 338

Db 270 TIGRTDFDGKLMAMSMAHPIDTPITGETFAFRYGFV-PPFLTYF FFDGSAKKQRDVPIFS 328

QY 339 LDQPTMMHDFAITENEVVVPDQQVVFK--LPEMTRGGSPVYDKNVARFGLDKYAED 395

Db 329 MTSPEFLHDATKRAIFAEQIQLGMRMNMIDLVLEGSPVGTDNGKTPRLGVIPKYAGD 388

QY 396 SSNIKWIADPCCPHLNWAESEPTTDEVVWIGSMNTPPSISNESDENIKSVSEIRLN 455

Db 389 ESEMKNWEEPGENI HAINAWDDEDGNSVVLAPIMSEHTLERMD-LVHAYEVKV KID 447

QY 456 LKGTGESTRRP LIISNEDQQVNL EAGMVNRNMLG RKTKEAYLALAEPWPKVSGFAKVDLTG 515

Db 448 LVGTGTVRRHPSIAR --- -NIDFAVINPAFLGRCSRYVAAAGGMPMKTSGYVLDVSKG 502

QY 516 E----VKKHYLGDNRYGGPLEFEGG---EDEGYLCLCFYHDEKTIWKSLEQIVNAV 566

Db 503 DRDDCTVAREMYSSCYGGEPFAYDPSNPEAEDDGIVVVTYVHDEVGTEKSFPLVMDAK 562

QY 567 S---LEVETVKLPSRVVPHGTFGADDLAK 596

Db 563 SPELEIVAVRPLPRVPRVPGFHGLFVYESDLNK 594

XX	RESULT 15
ID	AAG31334 standard; Protein; 517 AA.
XX	AC AAG31334;
XX	DT 17-OCT-2000 (first entry)
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 37613.
DE	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
XX	OS Arabidopsis thaliana.
XX	PR EP1033405-A2.
PD	06-SEP-2000.
XX	PR 25-FEB-2000; 2000EP-0301439.
XX	PR 25-FEB-1999; 99US-01211825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	23-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	28-APR-1999; 99US-0130891.
PR	30-APR-1999; 99US-0132048.
PR	04-MAY-1999; 99US-0132407.
PR	06-MAY-1999; 99US-0132485.
PR	07-MAY-1999; 99US-0132487.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	21-MAY-1999; 99US-0135533.
PR	14-MAY-1999; 99US-0134221.
PR	25-MAY-1999; 99US-0132486.
PR	07-MAY-1999; 99US-0132487.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-013524.
PR	21-MAY-1999; 99US-013553.
PR	24-MAY-1999; 99US-0135679.
PR	25-MAY-1999; 99US-0136221.
PR	27-MAY-1999; 99US-0136312.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-013722.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137702.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0138119.
PR	16-JUN-1999; 99US-0138452.
PR	17-JUN-1999; 99US-0138492.
PR	18-JUN-1999; 99US-0138454.
PR	18-JUN-1999; 99US-0138455.
PR	18-JUN-1999; 99US-0138456.

PR	18-JUN-1997	99US-0139457
PR	18-JUN-1997	99US-0139458
PR	18-JUN-1997	99US-0139459
PR	18-JUN-1997	99US-0139460
PR	18-JUN-1997	99US-0139461
PR	18-JUN-1997	99US-0139462
PR	18-JUN-1997	99US-0139463
PR	18-JUN-1997	99US-0139750
PR	18-JUN-1997	99US-0139763
PR	21-JUN-1997	99US-0139817
PR	22-JUN-1997	99US-0139839
PR	23-JUN-1997	99US-0140352
PR	23-JUN-1997	99US-0140354
PR	24-JUN-1997	99US-0140695
PR	28-JUN-1997	99US-0140823
PR	29-JUN-1997	99US-0140991
PR	30-JUN-1997	99US-0141827
PR	01-JUL-1997	99US-0141842
PR	01-JUL-1997	99US-0142154
PR	02-JUL-1997	99US-0142055
PR	06-JUL-1997	99US-0142324
PR	08-JUL-1997	99US-0142803
PR	09-JUL-1997	99US-0142920
PR	12-JUL-1997	99US-0142970
PR	13-JUL-1997	99US-0143542
PR	14-JUL-1997	99US-0143624
PR	15-JUL-1997	99US-0144005
PR	16-JUL-1997	99US-0144085
PR	16-JUL-1997	99US-0144086
PR	19-JUL-1997	99US-0144325
PR	19-JUL-1997	99US-0144331
PR	19-JUL-1997	99US-0144332
PR	19-JUL-1997	99US-0144333
PR	19-JUL-1997	99US-0144334
PR	19-JUL-1997	99US-0144335
PR	20-JUL-1997	99US-0144352
PR	20-JUL-1997	99US-0144632
PR	21-JUL-1997	99US-0144884
PR	21-JUL-1997	99US-0144814
PR	21-JUL-1997	99US-0145086
PR	21-JUL-1997	99US-0145088
PR	22-JUL-1997	99US-0145085
PR	22-JUL-1997	99US-0145087
PR	22-JUL-1997	99US-0145089
PR	22-JUL-1997	99US-0145192
PR	23-JUL-1997	99US-0145145
PR	23-JUL-1997	99US-0145218
PR	23-JUL-1997	99US-0145224
PR	26-JUL-1997	99US-0145276
PR	27-JUL-1997	99US-0145913
PR	27-JUL-1997	99US-0145918
PR	27-JUL-1997	99US-0145919
PR	28-JUL-1997	99US-0145919
PR	02-AUG-1997	99US-0145951
PR	05-AUG-1997	99US-0146386
PR	02-AUG-1997	99US-0146388
PR	03-AUG-1997	99US-0146389
PR	06-AUG-1997	99US-0147038
PR	04-AUG-1997	99US-0147204
PR	04-AUG-1997	99US-0147302
PR	05-AUG-1997	99US-0147192
PR	10-AUG-1997	99US-0148179
PR	11-AUG-1997	99US-0148319
PR	12-AUG-1997	99US-0148341
PR	13-AUG-1997	99US-0148565
PR	13-AUG-1997	99US-0148684
PR	16-AUG-1997	99US-0149368
PR	17-AUG-1997	99US-0149475
PR	18-AUG-1997	99US-0149476

PR	20-AUG-1999;	990S-0149723;
PR	20-AUG-1999;	990S-0149723;
PR	23-AUG-1999;	990S-0149902;
PR	23-AUG-1999;	990S-0149902;
PR	25-AUG-1999;	990S-0150566;
PR	26-AUG-1999;	990S-0150884;
PR	27-AUG-1999;	990S-0151065;
PR	10-SEP-1999;	990S-0153070;
PR	13-SEP-1999;	990S-0153758;
PR	15-SEP-1999;	990S-0154018;
PR	16-SEP-1999;	990S-0154039;
PR	20-SEP-1999;	990S-0154779;
PR	22-SEP-1999;	990S-0155119;
PR	23-SEP-1999;	990S-0155486;
PR	24-SEP-1999;	990S-0155659;
PR	28-SEP-1999;	990S-0156458;
PR	29-SEP-1999;	990S-0156596;
PR	04-OCT-1999;	990S-0157117;
PR	05-OCT-1999;	990S-0157733;
PR	06-OCT-1999;	990S-0157855;
PR	07-OCT-1999;	990S-0158039;
PR	08-OCT-1999;	990S-0158232;
PR	12-OCT-1999;	990S-0158389;
PR	13-OCT-1999;	990S-0159293;
PR	13-OCT-1999;	990S-0159294;
PR	18-OCT-1999;	990S-0159584;
PR	21-OCT-1999;	990S-0160741;
PR	21-OCT-1999;	990S-0160768;
PR	21-OCT-1999;	990S-0160770;
PR	14-OCT-1999;	990S-0159637;
PR	14-OCT-1999;	990S-0159638;
PR	14-OCT-1999;	990S-0159639;
PR	14-OCT-1999;	990S-0159640;
PR	25-OCT-1999;	990S-0161405;
PR	25-OCT-1999;	990S-0160844;
PR	21-OCT-1999;	990S-0160817;
PR	22-OCT-1999;	990S-0160980;
PR	22-OCT-1999;	990S-0160981;
PR	22-OCT-1999;	990S-0160982;
PR	26-OCT-1999;	990S-0161360;
PR	26-OCT-1999;	990S-0161361;
PR	28-OCT-1999;	990S-0161920;
PR	28-OCT-1999;	990S-0161992;
PR	28-OCT-1999;	990S-0161993;
PR	29-OCT-1999;	990S-0162142;
Query Match		30.4%;
Best Local Similarity		39.2%;
Matches 197; Conservative		9
Qy	112	PLPKTADPSVQIAGNFAPVNQE;
Db	22	PSRSPVDPKHVLSDNFAFVLD;
Qy	171	FUDGDMHAWKFFEHGSASYACRF;
Db	82	FUDGDMHAWKITHNGKATLCSR;
Qy	230	YARAAAGTVDPALGTGANAGLV;

Tue May 20 09:35:02 2003

us-09-758-269-6.rag

Page 14

Db 142 AARVLTGQYNPVNGTGLANTSIAFFSNRLFAIGESDLPVAVRLESGDLETIGRYDFDGK 201
QY 290 LESTIMIAHPKVDPESGEFLFALSYDVSKPVLKYFRFSPDGTKSPDVET-QLDQFTMMHDF 348
Db 202 LAMSNTAHPKTDPTITGETAFARYGPV-PPFLTYFREDSAGKQKDVP-IFSMTSFSFLHDF 260
QY 349 ATENFVVVVPDQQYVFK--LPEMIRGGSPVVDKKNKVAREFGILDKYAEDSNIKWIDAP 405
Db 261 AITKRHAIAPIAQGMRMNMNLIVLEGGSVGTDNGKTTRLGVPKYADESENKKWFEPV 320
QY 406 DCFCEFHWNAAWEPEDEVVIVGSCMTPPDSIFNESDENIKSVLSEIRLNKATGESTRP 465
Db 321 GFMNTHAINAWDEDDGNSVYLIAFNIMSIEHTLERMD-LVHALVEKVKDLYVGIVVRHP 379
QY 466 LISKEQDVNVNLEAGMVNRNLMGRKTFKAYLALAEPWPKVSGFAKVDLITGE---VKHK 520
Db 380 ISAR----NLDAFVNPAFLGROSRYYYAIGDPMPKTSGVVKLDVSKGDRDDCTVARR 434
QY 521 LYGDNRYGGEPFLPGEgg---BEDEGYTLCFVHDEKTWKSELQIYNAVS--LEVEATV 574
Db 435 MYGSGCYGGEPFFYARDPONPNEAEEDDGIVVTVHDEVTGESKPLVMDAKSPELEIYAV 494
QY 575 KLPSPRVYPYGFHGTFITGADDLAK 596
Db 495 RLPRRVYPYGFHGTVKESDLNK 516

Search completed: May 19, 2003, 15:47:11
Job time : 45 secs

BEST AVAILABLE COPY

GenCore version 5.1.4-p5-4578
 Copyright (c) 1993 - 2003 Compugen Ltd.
 OM protein - protein search, using sw model.
 Run on: May 19, 2003, 15:46:15 ; Search time 29 Seconds
 (without alignments)
 607.736 Million cell updates/sec
 Title: US-09-758-269-6
 Perfect score: 3150
 Sequence: 1 MASPTATAAVSGRWLGNNH.....VPYGFHGTFIGADDLAKQVV 599
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 262574 seqs, 29422922 residues
 Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0⁸
 Maximum Match 100⁸
 Listing first 45 summaries

Database : Issued_Patents_AA:
 1: /cgn2_6/podata/1/iaaa/5A_COMB.pep:
 2: /cgn2_6/podata/1/iaaa/5B_COMB.pep:
 3: /cgn2_6/podata/1/iaaa/6A_COMB.pep:
 4: /cgn2_6/podata/1/iaaa/6B_COMB.pep:
 5: /cgn2_6/podata/1/iaaa/PCTNS_Corb.pep:
 6: /cgn2_6/podata/1/iaaa/backtles1.pep:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	210	6.7	533	4	US-09-385-259-2	Sequence 2, Appli
2	210	6.7	533	4	US-09-645-370-3	Sequence 2, Appli
3	209	6.6	533	1	US-08-488-309A-6	Sequence 6, Appli
4	123.5	3.9	1043	3	US-08-928-361B-30	Sequence 30, Appli
5	123.5	3.9	1721	3	US-08-700-651.5	Sequence 5, Appli
6	123.5	3.9	1721	3	US-08-928-361B-6	Sequence 6, Appli
7	116.5	3.7	1042	3	US-08-228-361B-11	Sequence 11, Appli
8	116.5	3.7	1837	3	US-08-928-361B-5	Sequence 5, Appli
9	106	3.4	412	4	US-09-161-69-77	Sequence 77, Appli
10	104.5	3.3	2172	1	US-08-611-107-31	Sequence 31, Appli
11	98	3.1	741	4	US-09-001-984C-106	Sequence 106, Appli
12	97.5	3.1	10182	4	US-09-134-001C-31.59	Sequence 3159, Appli
13	97	3.1	516	1	US-08-997-B22-4	Sequence 4, Appli
14	97	3.1	516	1	US-08-577-403-1	Sequence 4, Appli
15	96	3.0	1410	3	US-09-335-409-3	Sequence 3, Appli
16	96	3.0	1410	4	US-09-668-102-3	Sequence 3, Appli
17	96	3.0	1410	4	US-09-567-960-3	Sequence 3, Appli
18	96	3.0	1410	4	US-09-568-480-3	Sequence 3, Appli
19	96	3.0	1410	4	US-09-568-486-3	Sequence 3, Appli
20	96	3.0	1410	4	US-09-568-472-3	Sequence 3, Appli
21	96	3.0	1410	4	US-09-567-899-3	Sequence 18, Appli
22	95.5	3.0	1802	4	US-09-122-478-18	Sequence 5, Appli
23	94.5	3.0	620	4	US-09-000-145-5	Sequence 188, Appli
24	94	3.0	1471	3	US-08-755-187-18	Sequence 2, Appli
25	92.5	2.9	868	2	US-08-907-166-2	Sequence 23, Appli
26	92.5	2.9	1057	4	US-09-697-367-23	Sequence 3, Appli
27	92	2.9	205	4	US-09-385-259-3	Sequence 17, Appli

ALIGNMENTS

RESULT 1									
US-09-385-259-2	Sequence 2, Application	US/09385259							
	Patent No.	620114							
	GENERAL INFORMATION:								
	APPLICANT: Aguirre, Gustavo D.								
	APPLICANT: Acland, Gregory M.								
	APPLICANT: Ray, Kunal								
	TITLE OF INVENTION: IDENTIFICATION OF CONGENITAL STATIONARY NIGHT BLINDNESS								
	FILE REFERENCE: 19603/2481								
	CURRENT APPLICATION NUMBER: US/09/385,259								
	EARLIER APPLICATION NUMBER: 60/103,219								
	EARLIER FILING DATE: 1999-08-30								
	NUMBER OF SEQ ID NOS: 8								
	SEQ ID NO: 2								
	LENGTH: 533								
	TYPE: PRT								
	ORGANISM: Canis familiaris								
	US-09-385-259-2								

Query Match Similarity 6.7%	Score 210; DB 4; Length 533;
Best Local Similarity 19.9%	Pred. No. 5 4e-13;
Matches 113; Conservative 98; Mismatches 218; Indels 140; Gaps 25;	
Qy 120 SVQI--AGNAPAPNEQVRRNLP---VVGKLPDSIKVYVRNGANPLH--FPTGHH 169	
Db 2 SIQVEHPAGGYKKLEVEELSSPILAHVTGIPMLTGSLLRCGGFLFEVGSEPH-YH 59	
Qy 170 FFDGDGMVHYAKEEHSASASYACRFQTINRFQE-----RQLGPVFPKAIGEIGHT 221	
Db 60 LFDDGQPLLHKDEKEHVYHRRFDTDAYRAMTEKRITYTEGTCFAEDPPCKNI----115	
Qy 222 GIARMLFYARAAGIVDPAHGTVANAGILYFENGRLAMS-----GEFLALSYDVSKPKYFRRS 326	
Db 159 TLETQVQLCNYSVNGATAAPHILENDGTVNIGNCFKNSTAYNTKIPPLQADKD 218	
Qy 327 PDGKSPDVEIQLD----OPTMMDFATENFVVDQCVVFKLPEMIRGGS-----374	
Db 219 P-1SKSEVYVQFPCDRKPSYHVSFGLPNIVFVKINLKFLSSWLGANNK 276	
Qy 375 -----PVYDVKYVARGLDKYADESSNIKWIADPCFCFLWNAAWEPEPTDEV 424	
Db 277 DCFESNETMGWHLTADKKRKYLNKRNKRTSSFNL-----FHINTYDNEF-L 324	

Query 425 VVIGSCMTPPDSIFNESD-ENIKSVLSEIRNLKLTGESTRRPIISNEDQQVNLLEAGMVNR 483
 Database 325 IVDLCCWKGFEVVNYLNARENNEEVK--KNARKAPQPEVRSSVLPLNIDKADTGK 381
 Query 484 NMGLRKTKFAYAL-----AEPWKVSG-----FAVKDLTT-GEVKKHLYG---- 523
 Database 382 NVTLPNTTATLRSDETIWLEPEVLSGPQAFEFQIYQKGKPYTAYGGLNH 441
 Query 524 -----DNRYGGEPEPLP-GGGGEEDGYILCFVHDEKTWK--SE 559
 Database 442 FVPDFRCLKNVTKTETWWQEDPSYSISEPITVSHDALEDDGVVLSVVSPGAGKPAY 501
 Query 560 LOTINAVSLEVEATVLPSPRVYGFHTF 588
 Database 502 LLILNAKDLSEVARAEVEINIPVTEHGF 530

RESULT 3
 US-08-488-305A-6
 Sequence 6, Application US/08488305A
 ; Patent No. 567972

GENERAL INFORMATION:
 ; APPLICANT: Bvik, Claes Olof; Eriksson, Ulf; Peterson, Per A.
 ; TITLE OF INVENTION: Isolated Protein Receptors, Antibodies Which
 ; bind Thereto, Nucleic Acid Sequence Coding
 ; PATENT NO. 567972
 ; TITLE OF INVENTION: Therefor, And Uses Thereof
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felfe & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 1002

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC-DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,305A
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kohli, Vineet
 ; REGISTRATION NUMBER: 37,003
 ; REFERENCE/DOCKET NUMBER: LUD 5280.3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 688-9200
 ; TELEFAX: (212) 838-3884
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 533
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 US-08-488-305A-6

Query Match 6.78, Score 210, DB 4; Length 533;
 Best Local Similarity 19.98, Pred. No. 5.4e-13; Gaps 25;
 Matches 113; Conservative 98; Mismatches 218; Indels 140;

Query 120 SYQI---AGNEFAPVNEOPVRRNLP---VVGKLUDSTKGVYWRNGANPLH--EPYTGHH 169
 Database 2 S1QVEHAGGYKLFEELVEEVSPLTAHVGTIPLENLTGSURCGPGLFEVGSEPP-YH 59

Query 170 FEDGDGMHAYKVFEGSASYASRETQINRFRE-----RQLGRFVFPKAIGE, Length 221
 Database 60 LEDGOALHKFEKEGHVTVYHRFRFIRDAYRAMTEKRIVITEFGICAFDPCKNI --- 115

Query 222 GSTARMLFYARAAGAYIDPAHGTGVANAGLVYFNGRJLAMSD----DLIVQVOTTPNG 276
 Database 116 -RSFEEFYE-----REVEVDNALV---NVYPVGDYACTETNITK NPE- 158

Query 277 DLIKTVGRFDGQLE-STMIAPKVDPES-----GELPALSYDVSKPKLYFRFS 326
 Database 159 TLETIKQVDLCNVVSUNGATAPHIENDGTVVNIGNFGKNSIAYNIVKTEPLQADKED 218

Query 327 PDGTKSPLDVEICOLD----OPTMMHDFAITENFVYVEDQVVKLPEMIRGGS----- 374
 Database 219 P - ISKSEVVVOFCSDRFKPSVHSLPQIVYFETPVKINLKFLSSWLGANYM 276

Query 375 -----PVYDKNKVARFGLDKYAEDSSNWKIDAPDFCFHLNAMEEPEDEV 424
 Database 277 DCFFESNETMVGWIIHAKKKRKYLNKXRTSFNL-----FHINIVEDNEE-L 324

Query 425 VVIGSCMTPPDSIFNESD-ENIKSVLSEIRNLKLTGESTRRPIISNEDQQVNLLEAGMVNR 483'

Query Match 3.98; Score 123.5; DB 3; Length 1043;
 Best Local Similarity 20.6%; Pred. No. 0.0021; Mismatches 194; Indels 203; Gaps 35;
 Matches 125; Conservative 85;

Qy 5 TATAAVSGRWLGGNHTQPLLSLOSSDLSYCSSL-PMASRVT'RK---LNVSSALHTPP 58
 Db 153 TAGIIVSG-----ISASESL-LSQKSALIDPATNMVYGEFGGLNPATGVMTFG 200

Qy 59 ALHPKQSSNSPAT---VVKPKARESNPKOMNL-----FQRAAALD 98
 Db 201 FLGPSEQIQSPELEDDGILIPPEVAANADKEKLSSIPSPVSPESIPEKDOKIDSISELMYD 260

Qy 99 AAEGFLVSHEKHLPLKTADEPSVQIAGNAPVNEOPVRNLPVYKGKLPDSIKGYVVRNGA 158
 Db 261 LESGRJIGOVSKRPIFGS-----LAGDLNPIMKPTQTD-SVTKPDTPTGGL-----307

Qy 159 NPLHEPVGTGHFFDGDGMYHAYKEFHGSASY--ACRFTQNRFYQEROLGRPVFPKAIGE 216
 Db 308 -PFENP-TGH-----LINPTNNNTMDSSFAGAYKAVSNGIKTDNVYGLPV-----GE 353

Qy 217 LHGTGJARMLFYARAAGAVTDPHGTGYAN-AGLVYENGRLLAMSDDDLPYQVQIT 273
 Db 354 ITGLPKDPSDIPNSTTGELVDPSGKPINNSTAGIV--SGKPSYDPIEDE-----403

Qy 274 PNDLKRIVGRFD----FDGQLESTMIAPKVDPESELFAVSKPPLYKFRFSP 327
 Db 404 -NGNL----FDPSINLPIGN-----NQLVNPNETNSTVSGSTGSTKIP-----KP 443

Qy 328 ----DGTKSPDVEI-QLDO-----PTNMHDFTENFVVVPDQQVVF 364
 Db 444 GIPVNNGGVVDEEARDQKGDKGLIVPPINSINKDPVNNTQYSNTGNINPETGKV- 502

Qy 365 KLPENTRG-----GSPVYDKNKWAFRGFLDKFEDSSNIKWIDAPDCFCFHLMNAWE 417
 Db 503 -IPGSUPLSINYPSFTIPQQTDE-----ITGKPDVTVGLPY -----538

Qy 418 EPETDVEVVIGSCMTPDFNESDENLNKSYLSEFLNLKGETSERPISNEDQVNLE 477
 Db 539 DPSTGELIDPATKLKPFGSYGD----ELTEV.LNTTDEVIGLPI-----DLE 583

Qy 478 AGMVNRNMLGRKTKEFAYLALAEPWPKVSGFARK--DLTGEVKRHYGDNRGGPLFL 534
 Db 584 TGL-----PRDFVSGLPOLPNGLVDESNKKEIPGS-HSG---FI 619

Qy 535 PGEGGEE 541
 Db 620 NGTSGBQ 626

RESULT 5
 US-08-928-361B-30 ; Sequence 30, Application US/08928361B
 ; Patient No. 6071518
 ; GENERAL INFORMATION:
 ; APPLICANT: Petersen, Carolyn
 ; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 ; POLYMERS, POLYNUCLEOTIDES, PROTEINS, ANALOGS AND FRAGMENTS
 ; OF PROTEINS, POLYNUCLEOTIDES, POLYNUCLEOTIDE ANALOGS AND FRAGMENTS
 ; OF PROTEINS, FOR TREATMENT AND DIAGNOSIS OF CRYPTOSPORIDIUM
 ; SPECIES INFECTIONS
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; PETERS, VERNY, JONES & BIKSA
 ; STREET: 385 Sherman Avenue, Suite 6
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306-1840
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0. Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/328,361B
 ; FILING DATE: 12-SEP-1997
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/026,062
 ; FILING DATE: 13-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: VERNY, Hana
 ; REGISTRATION NUMBER: 30-518
 ; REFERENCE/DOCKET NUMBER: 480-76-1 (HV)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-324-1677
 ; TELEFAX: 650-324-1678
 ; INFORMATION FOR SEQ ID NO: 30:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1043 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Protein
 ; ORGANISM: Cryptosporidium parvum

US-08-700-651-5
 ; Sequence 5, Application US/08700651B
 ; Patent No. 6015882
 ; GENERAL INFORMATION:
 ; APPLICANT: PETERSEN, CAROLYN
 ; APPLICANT: LEITCH, JAMES
 ; APPLICANT: NELSON, RICHARD, C.
 ; APPLICANT: GUT, JIRI
 ; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
 ; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
 ; FILE REFERENCE: 460-19-4 (HV)
 ; CURRENT APPLICATION NUMBER: US/08/700,651B
 ; CURRENT FILING DATE: 1997-08-14
 ; EARLIER APPLICATION NUMBER: 08/415,751
 ; EARLIER FILING DATE: 1995-04-03
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1721
 ; TYPE: PRT
 ; ORGANISM: Cryptosporidium parvum

US-08-700-651-5

Query Match Score 3.9%; Length 1721; DB 3; Best Local Similarity 20.6%; Pred. No. 0.005; Matches 125; Conservative 85; Mismatches 194; Indels 203; Gaps 35;

QY 5 TATIANSGRMLGNGNHTOPPLSSQSDSLSCSSL - PMASRVRTRK --- LNVSSALHTPP 58
Db 831 TLAGIVSG --- ISASESL - LSOKSALLDPATNMVVGEGGLINPATGYMIPG 878

QY 59 ALHIFPKOSSNSPAI --- VVKPRAKESNTKOML --- FQRAAAAALD 98
Db 879 FGQSEQTQESPEEDEGGGLIPPEAAANADKFKLSTIPSPVSEPEKDKDITSELMYD 938

QY 99 AAEGLFLYSHEKHLPLKPTADPSVQIAGNFAPVNEOPVVRNLPVYKGKLPSIKGYVVRNGA 158
Db 939 IESGRLLIGQSKRPIGS --- TAGDLNPIMKTPOTD - SVIGKPTLPTTGL --- 985

QY 159 NPLHEPVTHGHHFEDGDMYHAKVKEHGSASY - ACRFQTQENRVQEROLGRPFVKPAIGE 216
Db 986 -PENAPP-TGH--- -LNINPTNNNTMDSSFEAGAYKA-SNGIKTDNVYGLPV---GE 1031

QY 217 LIGHITGIARMLFYARAAGIVDAHGTGAYN - AGLVYFNGRL-LAMSEDDLPYQVQIT 273
Db 1032 ITGLPKDPKSDIDPENSTIGELVDSTGKPTINNSTAGIV - SGK6SLPPTEDE --- 1081

QY 274 PNDGLIKTVGRFD --- FDGOLESTMIAHKPVDPESGEFLAFLSYDVSKPVLKYFRFSP 327
Db 1082 -NGNL--- -FDPSTNLPLDGN --- NQLVNPETNSTVSGSTS GTKPK --- KP 1121

QY 328 ----DGTRSPDVEI - QLDQ --- PIMMHDRATIENFVVVDQVVF 364
Db 1122 GIPVNGGNGVPDEPAKQDKGKDGGLIVPPTNINSINKDPVINTYSNTGNIINETGKV- 1180

QY 365 KLPDEMIRG --- GSPVYWDKANKVAREGILDKYAEODSSNIKWIDAPDCFCFHWNANE 417
Db 1181 -IPSLSLPGSLNPPSENTPQTDE --- ITGPVDTVYGLPY --- 1216

QY 418 EPFDEVVVGSCTMPDSEFNESDENLKSVLSEFIRLNLTGTGESTTRRPTISNEDQVNLE 477
Db 1217 DPTGEIIDPATKLPKIPGVAGD --- EILTEV-LNITDTEVTLGLP1 --- DLE 1261

QY 478 AGMVNRNMLGRKTKFAYALAEPPVKVSGFAKV -- DLTTGEVKKKHLYDNRVKGGEPLPL 534
Db 1262 TGL --- PRDPVSSLPQLPQNLPTVDPNSNKKPIFGS - HSG -- FTI 1297

QY 535 PGEGEE 541
Db 1298 NGTSSEQ 1304

RESULT 6

US-08-928-361-B⁶

Sequence 6, Application US/08928361B
Patent No. 6071518

GENERAL INFORMATION

APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS, THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS OF INVENTION: THEIR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
NUMBER OF SEQUENCES: 30

TITLE OF INVENTION: SPECIES INFECTIONS

CORRESPONDENCE ADDRESS:

ADDRESSEE: PETERS, VERNY JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840

COMPUTER READABLE FORM:

MEDIUM: Floppy disk
COMPUTER: IBM PC compatible

QY 535 PGEGEE 541
Db 1298 NGTSSEQ 1304

QY 535 PGEGEE 541
Db 1298 NGTSSEQ 1304

RESULT 7
US-08-928-361B-11
Sequence 11; Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentnet Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/025,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verry, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480-76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1042 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-361B-11

Query Match Score 3.7%; Score 116.5; DB 3; Length 1042;
Best Local Similarity 19.9%; Fred. No. 0.011;
Matches 113; Conservative 74; Mismatches 175; Indels 207; Gaps 30;

QY 48 LNVSALIHTPDLHFKPKQSSNSPAI---VVKPKAKAESNTKQMLN-----88
Db 190 LNPATASYMIPSLGPSEQTPPSPEIPEGGIPPEPEVAANAKDFKLSSPPSYPEKQD 249
QY 89 -FQRAAAALLDAEAEGLFLVSHBKLHPKTAPEPSVQJAGNEAPVNEOPVRNLPVYGSKLPD 147
Db 250 KIDSISELMYDIESGLIGQYSKRPIGS-----TAGDNPIMPTQTQ-SVYGKPID 302
QY 148 SIKGYYVNRNGANPLHEPVTGHFFDGMVYIAVKFHGSASY-ACRFETONRFYQERQL 205
Db 303 EITGTL-----PENNP-TG-----LINTPNNTMDSSFAGAYKYAANGIKTDNVY 347
QY 206 GRPV-----FRKAIGELHAGTGIARLMLFYRAAAGIVDEFAHGTGVAN-- 248
Db 348 GLPVDEITGLPKDPVSDFPNSTTGEL-----VDFSTGKPNNYT 387
QY 249 AGLVVNGRLIAAMESDDLPYQVQTPNQGDL---KTYGRFDFGOLESTMIAHAKPYDPESG 305
Db 388 AGIV-SGKRGKPLPIED-----NGNLIDPSKHLPLIBSNQNQ-----VNPDEN 428
QY 306 ELFALSVDYVSKYLYKFRESP-----DGTKSPDVEI---QLDQ-----VQYB; 342
Db 429 STVSGSTSGSKP-----RPGIPVNGGEVVPDDEAKDQADKGKDGLTVPTNSINKDP 481
QY 89 -FQRAAAALDAEGLFLVSHEKLPKTADPSVQJAGNFAPVNEQPVRRNLPVVKLKD 147

RESULT 8
US-08-928-361B-5
; Sequence 5; Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 9306-1840
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/025,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verry, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480-76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1042 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-361B-5

Query Match Score 3.7%; Score 116.5; DB 3; Length 1837;
Best Local Similarity 19.9%; Pred. No. 0.03;
Matches 113; Conservative 74; Mismatches 175; Indels 207; Gaps 30;

QY 48 LNVSALIHTPDLHFKPKQSSNSPAI---VVKPKAKAESNTKQMLN-----88
Db 95 LNPATVMIPGSLGP-EQTFSPETEDGGIPIPPEVAAANAKDFKLSSPPSYPEKQD 1044
QY 89 -FQRAAAALDAEGLFLVSHEKLPKTADPSVQJAGNFAPVNEQPVRRNLPVVKLKD 147

RESULT 9
OS-09-461-607-77
; Sequence No. 6277974
; GENERAL INFORMATION
; APPLICANT: COGEN NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stewart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSTICING,
; TITLE OF INVENTION AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-993
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 77
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
; OS-09-461-697-77

Query Match 3.4%; Score 106; DB 4; Length 412;
Best Local Similarity 20.6%; Pred. No. 0.027; Mismatches 131; Indels 110; Gaps 16;

Matches 74; Conservative 45; Mismatches 131; Indels 110; Gaps 16;

Qy 23 PLSSGSDLSYCSSLPMASRVTRKLNVSSAHLTPALHFKQSSNSPAIVKPKAK-- 79
Db 9 PISSSQTNRSSP-DLLPREFRIVE-----VHDPP-LHOPSANKPKPTMLDIPSEPCS 59

Qy 80 -----ESTNTKOMNLFORAAAALDAEGLFYSHEKELPLPKTADPSVQIAGNFAPY 130;
Db 60 LIHTIQLIGHNRRLNLIATAQAQNQQTEG--VKESESEPLP-SCPGSPPLPDLLPL 116

RESULT 10
US-08-611-107-31
; Sequence 31 Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,107
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US SN 07/956,700
; FILING DATE: 02-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US93/09340
; FILING DATE: 30-SEP-1993
; PRIORITY DOCUMENT NUMBER: ARCD:221
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOMETRY: linear
; US-08-611-107-31

Query Match 3.3%; Score 104.5; DB 1; Length 2172;
Best Local Similarity 19.9%; Pred. No. 0.74; Mismatches 80; Indels 147; Gaps 27;

Qy	66	SSNSPAVIVVKPAKESNTKQMNLFQRAAAALDAABSFVLYHEKHLPLPKTADPSVQTAG 125	Db	138	LYDALYGTGTDVLPETDGAEKGPTNYKVRGDKVLIAYARKFLD-----DSVPLOSSGSFGDATG 192
Db	1406	TGNTCTYDIVEYEDNTNH--LEYSATPTAGPLH--LAHEPYKPL---DAIDLKR 1457	Qy	193	FTQTNRFVQEROLGRPVPEPKA----GELIGHTYAR----LMFYARAAGTVDPAH 242
Qy	126	NEFAPYEQVRRNLPVVGK-----PDSIAGVYVRNGANLPHETPGTHFFDGDIM-- 176	Db	193	FT-----VQDGQLVVAPDKSTGLANGQFAGTYGAESSPSYVLLNLHTEILDPE 247
Db	1458	AAARKNETTYCYDFPLAFETALKKSWSGGISHVAESNEHHORYAEVTTELIAADSTSWGT 1517	Qy	243	GTGVAN-AG-----LIVYFNGRLLAMSDD--LPYQVOTTPN-GDL-KTV---- 281
Qy	177	-VHAKEEHGSASYACRFTQTNRFVQEROLGRPVF-----PKAIGEHHGF 221	Db	248	QVGTIDRAGVKDVTLLESAITTMDPDSVAAYDVKDG 307
Db	1518	PLVPVERPPGSNRFGVYAWNMKLSTPEFPGRREITVYANDVTFKAGSFGPREDADFDAYT 1577	Qy	282	-----GRFDEFQG-----ESTMIAHKVDPDESGELFALSVDVVS 316
Qy	222	GJA--RLMLYYARAAGTIVDPAHTGVAANGLVYNGRILLAMSEDDLP----YQVQITP 274	Db	308	TAFLRVLNDRDRNYPAGGGQFTLPGRSUMLFYRNVGHLMTNDIAVDTGSEPFEGIMDHALF 367
Db	1578	NLACERKIPPLYLSSATAGA----RGVAAEIKACH--VGWSDQSPERGEHYIYLIE 1629	Qy	317	KYLYKFERFSPDGTQSPBPDV-----ETQDQPTMMH--DEFAIT-ENFVAVPBDQQV 362
Qy	275	NGDLKTVGRFDEDQGLESTMIAHPKVDPDESGELFALSVDYVSK----- 317	Db	368	TGLI-----AINGLKASDVGPLINSRTGSIYIKPV KMHGDAEVFTCELSRVED-- 418
Db	1630	Q-----DY-SRLLLSSVTAHELKVPESSETRWWVDTIYGKEDSLGCENLHGSSAIAAS 1679	Qy	363	VFKLPEMTRGSPPVYDKNRYARFGILDKYAEQDSNNIKWIDAPDCFCFHWNNAWEPEPTD 422
Qy	318	PYLYKFR--FSPDGTKSPDVEI-----QLDOPTMMHDFAITENFVYVPDQOV 362	Db	419	VLGLPQ-----NFMKIGMDEERTTIVNLR-----ACTRAAAD 451
Db	1680	AYSKAYRETFTILEVTGTRAIIGAYLARLGMRCTIQRLDQPILTGSYALNKL--GREY 1736	Qy	423	EYVVTGSCMTTPDSIFNESDENLKVSLSEIRUNLKTESTRPP-----LISNEQQVN 475
Qy	363	VFKLPEMTRGSPPVYDKNRYARFGILDKYAEQDSNNIKWIDAPDCFCFHWNNAWEPEPTD 422	Db	452	RVFINIGF-----LDRTGDELTISMEGPMVAKGTMKSOPWLAYERHNVD 498
Db	1737	YSSQML--OGPKIMATNGVYHLTYSVSDLEGVSALKWL-----SYVZPYVG 1781	Qy	476	-----KTKFAY--LALAFPW-PKVSG-----FAKY 510
Qy	423	EYVVTGSCMTTPDSIFNESDENLKVSLSEIRUNLKTESTRPPITSNEDQOQNLDAGMVN 482	Db	499	AGLAAGFSGRAQVQKGKGMWTTFELMDMVETKAQPRAGASTRAWPSPTAATLHALFYHQV 558
Db	1782	GPLPIVKSLLDPERAVTYFPEN-----SCDAAARACICQDQGKWLGMED 1827	Qy	511	DLTGEGVKKHLYGDNRYGGEPL-----FLPEGGEDEDE-----GYILCFYHDE 553
Qy	483	RNMLGRKTKFAYLALAEPWPK--VSGEFAKYVDTTGEVKHLYGDNRYGGEPLELPG----- 536	Db	559	DVA--VQGQLAGKRATIEQLLTIPIAKEJAWAPPDIREEVDNNCQSILQYVVRVV-DQ 615
Db	1828	RE-----SFVETLEGWAKTIVGRAKSAAKTAQA--LLDFNR-EELPLEFLANWRQ 1875	Qy	554	KTWKSELQIVNAYSL-EVEATYKLPPSRVYPGF--HGTFIGAD 592
Qy	537	-EGGEED--EGYI--LCFVHDEKWKSE--LQIVNAVSELEVEATYKLPPRV 580	Db	616	GVGCSKVPDHDVALMEDRATLRISSOLLANLHRGVITSAD 657
Db	1876	FSGQQRDLFEGTQLAGXMTYKQPAFVYTPKAGELRGGANWVVDTSKI 1927			RESULT 12 US-09-134-001C-3159
					Sequence 3159, Application US-09134001C ; Patent No. 6380370
					GENERAL INFORMATION: ; APPLICANT: Lynn Doucette-Stamm et al ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAA ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS ; FILE REFERENCE: GTC-007
					; CURRENT APPLICATION NUMBER: US/09/134,001C ; CURRENT FILING DATE: 1998-08-13 ; PRIOR APPLICATION NUMBER: US 60/064,964 ; PRIOR FILING DATE: 1997-11-08 ; PRIOR APPLICATION NUMBER: US 60/055,779 ; PRIOR FILING DATE: 1997-08-14 ; NUMBER OF SEQ ID NOS: 5674 ; SEQ ID NO 3159 ; LENGTH: 10182 ; TYPE: PRT ; ORGANISM: Staphylococcus epidermidis ; US-09-134-001C-3159
					Query Match 3.1%; Score 98; DB 4; Length 741; Best Local Similarity 21.7%; Pred. No. 0.53; Gaps 37;
					Query Match 3.1%; Score 97.5; DB 4; Length 10182; Best Local Similarity 19.4%; Pred. No. 6.1; Mismatches 204; Indels 165; Gaps 66
					Query 13 RWLGGMHTQPLLSSSSSDSYCYSPPSPMASTVRL-----NVSSALHPPALIFPKQS 66 ; Db 673 RWQGAN-----AMNAQIEELFSEQENIQAARSGRP1QFLGVGFDVEDSHHNPETL-LPVN- 726 ; Qy 67 SNSPATIVPKAK-----ESNKKONLFLQRAAAALDAARGFLYSHESKHLPLPKT 116 ; Db 702

Query 117 ADPSVQAGNFAPVNEQPVRRNLPVVGKLPDSI--KGVYVRNGANPLHEPVYGHHFEDG 173
 Database 782 OEPRIKTSN----EPIKDSWSTIYTGPNUMLQAVGRNNTATEKLNLYGH-IDP 833
 Query 174 DGMVHAYKFEGHSASYAARFTQINRFVQERQJGRPVFKPAI--GEIYGH----T 221
 Database 834 GNYFITVKGDKYEQFEIR----SKPTPRITITANERLGNPNHKPEIRVT 880
 Query 222 GIALMLFYARAAGIYD-----PAHGTYVANAGLYF----MGRLLAME 263
 Database 881 DIFNDTPAKIKLYMGGTDHDPDEINPYTPEVVAEA--YHDNDPSKNGVLTERRS 937
 Query 264 D---DLPYQVQITPQNGDLKTYGRF----DFDGQLE-STMAAHPKVDPESSGELFALSYD 313
 Database 938 DYLKDPL----SGEJKAIYQQYQVSNNNSNSVPFSDDITPPTINEPABLVH---- 986
 Query 314 VYSKPQLYKF-----FSPDGTKSPDVEIQLDQPTNMHDFATTENVVPPDQV 362
 Database 987 -----KYYRGDHVEITLPTVDTNTGGSLRLDVNVNLPO-----GWTKTFLINPNNT 1032
 Query 363 VFLKPEMIRGGSPVYDKRNKAFLGILDKYAEADSSNIKWIADPDCFCFHLYNAWEET- 421
 Database 1033 EGTL-KLIGNITSNEAQXNTTHFNIT--ADTNSG-----NNTNPATF 1072
 Query 422 -----DEVWYTG-----SCMTPPDSEFINESDENLSKVSEIRLNKTEGESTRPII 467
 Database 1073 IIAVNGKLAADDLNPGVLSSDQLQVLTDPSSLNSEREEVYKRKSEANANIRSYLONNPIL 1132
 Query 468 SNEDQQV 474
 Database 1133 AGYNGDV 1139

RESULT 13
 US-08-097-829-4
 ; Sequence 4, Application US/08097829
 ; Patent No. 5498831
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgess, Diane G.
 ; ATTORNEY: Dooner, Hugo K.
 ; TITLE OF INVENTION: Pea App-Glucose Pyrophosphorylase Subunit
 ; TITLE OF INVENTION: Genes and Their Uses
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Khourie and Crew
 ; STREET: One Market Plaza, Steuart Tower, Suite 2000
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/097,829
 ; FILING DATE: 19930723
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heslin, James M.
 ; REGISTRATION NUMBER: 29,541
 ; REFERENCE/DOCKET NUMBER: 12176-35
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-326-2400
 ; TELEFAX: 415-326-2422
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 516 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-097-829-4

Query Match 3.1%; Score 97; DB 1; Length 516;
 Best local Similariy 20.9%; Pred. No. 0.36; Mismatches 172; Indels 224; Gaps 32;

Query 1 MASFTATAVSGRGLGGNHTOPPLSSSSDLSYCSSLPMASRVTRKLNVSSA----L 54
 Database 1 MASMAAIGCVLK----SSSSSKAARNLSETSQLCGDKI 46
 Query 55 HTPPALHPKQSSNSPAIVVKPA-KESNTKOMNLQAAALDAEAGFLV--SHEKL 110
 Database 47 FT--VSSTRSSSGRNPTVSKAVSDSKNSQTCLDPASRSV--GIUUGGACTRL 99
 Query 111 HPL-PKTADEPSVATGNAFPVNEQPVRRNLPVVGKPLPSKIGVYRNGANPLHEPVYGH 169
 Database 100 YPLTKRKPKPAVFLGANVRL-----DIPVSNCNLNSNISKVLTQFN--SASENRH 149
 Query 170 F-----FDGDMYTHAVKEE-----HGSA---SYACRETQTN-----197
 Database 150 LSRAYASNLGGTKNEGFVEVLAQQSPBPNNPWFQGTADAVRYLNLFEEHNYLEYLYLAG 209
 Query 198 ---RFVYQERQLQRPVFFKAIGELHGHTGJIARLMFLPARAAMGIVDPAHGTGVANAGLV 253
 Database 210 DHLYRMDYER-----FTQAHRESDADTVASLPMDERA-----FOLMK 250
 Query 254 FN-GRLIAMSE-----DD-----LPYQVQITBNG-----D 277
 Database 251 IDEGRIVEESEPKGEOLSTMIAHPKVDEPSEGEL--FALSYD-----313
 Query 278 L---KTYGRFDFGQLESTMIAHPKVDEPSEGEL-----313
 Database 308 LLRDKFPPANDFGE-----VFGATEGLRQVAYLYDGYWEDIGTLEAFYNANL 357
 Query 314 -VYSKP-----YLYKFERFSIDGTRSPDVETQDFOFTMMIDFATTENFV---VYFDPDQV---362
 Database 358 GITKKPVDSEFDRSSPDIYTOP-----RYLPSPKMDADITDVSIGCIVKNCKHHS 412
 Query 363 VFKLPEM-----RGSPPVYDKNKVARTGILDKYAE 394
 Database 413 VVGLRSCSEGALIEDTLLMGADYETTAADRRELAAGGVPIGIGKNSHIKRAITDKNAR 472
 Query 395 DSSNIKWDADPOFCFHWNWEEPE----TDEVVVIGSCMTTPDSTI 437
 Database 473 IGDVKLINS-----NYQEARAFETEGYEIKSGTIVYKDALIPSGIV 515

RESULT 14
 US-08-577-403-4
 ; Sequence 4, Application US/08577403
 ; Patent No. 5773693
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgess, Diane G.
 ; ATTORNEY: Dooner, Hugo K.
 ; TITLE OF INVENTION: Pea ADP-Glucose Pyrophosphorylase Subunit
 ; TITLE OF INVENTION: Genes and Their Uses
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Khourie and Crew
 ; STREET: One Market Plaza, Steuart Tower, Suite 2000
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/097,829
 ; FILING DATE: 19930723
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heslin, James M.
 ; REGISTRATION NUMBER: 29,541
 ; REFERENCE/DOCKET NUMBER: 12176-35
 ; TELEPHONE: 415-326-2400
 ; TELEFAX: 415-326-2422
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 516 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-097-829-4

PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/577,403
 ; FILING DATE: 22-DEC-1995
 ; CLASSIFICATION: 800
 ; PRIORITY APPLICATION DATA:

RESULT 15
US-09-335-409-3
; Sequence 3, Application US/09335409
; Patent No. 612129
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Liwon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Ziffle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Clegg, Thomas
; APPLICANT: Tamm, Michael



GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 15:48:44 ; Search time 58 Seconds
(without alignments)
996.104 Million cell updates/sec

Title: US-09-758-269-6
Perfect score: 3150
Sequence: 1 MASFTATAAVSGGRWLGGNHT.....VPYGFHGTFIGADDIAKQQV 5 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:
 1: /cgn2_6/ptodata/2/pubpaas/US08_NEW_PUB.pep:
 2: /cgn2_6/ptodata/2/pubpaas/PTCT_NEW_PUB.pep:
 3: /cgn2_6/ptodata/2/pubpaas/US06_NEW_PUB.pep:
 4: /cgn2_6/ptodata/2/pubpaas/US05_NEW_PUB.pep:
 5: /cgn2_6/ptodata/2/pubpaas/US07_NEW_PUB.pep:
 6: /cgn2_6/ptodata/2/pubpaas/US07_PUBCOMB.pep:
 7: /cgn2_6/ptodata/2/pubpaas/PTCTUS_PUBCOMB.pep:
 8: /cgn2_6/ptodata/2/pubpaas/PUBCOMB.pep:
 9: /cgn2_6/ptodata/2/pubpaas/US09_NEW_PUB.pep:
 10: /cgn2_6/ptodata/2/pubpaas/US09_PUBCOMB.pep:
 11: /cgn2_6/ptodata/2/pubpaas/US10_NEW_PUB.pep:
 12: /cgn2_6/ptodata/2/pubpaas/US10_PUBCOMB.pep:
 13: /cgn2_6/ptodata/2/pubpaas/US60_NEW_PUB.pep:
 14: /cgn2_6/ptodata/2/pubpaas/US60_PUBCOMB.pep:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3150	100.0	599	10	US-09-758-269-6
2	2280.5	72.4	605	10	US-09-758-269-16
3	2168.5	68.8	612	10	US-09-758-269-12
4	1939.0	63.2	583	10	US-09-758-269-2
5	1663.5	61.3	604	10	US-09-758-269-14
6	968	52.8	577	10	US-09-758-269-10
7	938	30.7	595	10	US-09-758-269-4
8	937	29.8	538	10	US-09-758-269-8
9	937	29.7	538	10	US-09-758-269-18
10	936	13.1	505	1	US-09-758-269-33
11	411.5	8.3	526	9	US-10-053-192-1
12	260	7.8	506	9	US-10-053-192-4
13	185.5	5.9	529	9	US-10-053-192-5
14	107	3.4	26926	9	US-09-759-508B-2
15	106.5	3.4	2701	9	US-10-171-261-83
16	106	3.4	412	10	US-09-922-261-77
17	105	3.3	5701	10	US-09-864-761-37319
18	103.5	3.3	1311	9	US-10-103-377C-6

8

Result No.: 6
Length: 599
Type: PRT
Organism: Arabidopsis thaliana

Query Match 1 US-09-758-269-6
Best Local Similarity 100.0%; Score 3150;
Matches 599; Conservative 0; Indels 0; Gaps 0;

QY 1 MASPTATAAVSGRNIGNGNTOPPLSSQSDSLSSYCSLPMASRVTRKINVSSALHTPPAL 50
1 MASPTATAAVSGRNIGNGNTOPPLSSQSDSLSSYCSLPMASRVTRKINVSSAHTTPAL 60
Db 1 HFPKQSSNSPAIVVKPKAKESNTQKOMLFQRAAAALDAAEGLVSHESKLHPPLKTAQPS 120
1 MASPTATAAVSGRNIGNGNTOPPLSSQSDSLSSYCSLPMASRVTRKINVSSAHTTPAL 60
QY 61 HFPKQSSNSPAIVVKPKAKESNTQKOMLFQRAAAALDAAEGLVSHESKLHPPLKTAQPS 120
Db 61 HFPKQSSNSPAIVVKPKAKESNTQKOMLFQRAAAALDAAEGLVSHESKLHPPLKTAQPS 120
QY 121 VQIAGNAPVNQPVRLPVGKLPDSIKGVYRNGANPLHEPVTHFFGDGMSYHAY 180
Db 121 VQIAGNAPVNQPVRLPVGKLPDSIKGVYRNGANPLHEPVTHFFGDGMSYHAY 180
QY 181 KFENGSAVASACRFQTQNREVQERQLGRPVFKAIAGELHGHGTGIAIRMLFYARAAGIVYDP 240

ALIGNMENTS

Sequence 13187, A
Sequence 5107, Appl
Sequence 54, Appl
Sequence 21, Appl
Sequence 21, Appl
Sequence 58, Appl
Sequence 3692, Appl
Sequence 5166, Appl
Sequence 10929, A
Sequence 3, Appl
Sequence 1056, A
Sequence 18, Appl
Sequence 2, Appl
Sequence 1332, Appl
Sequence 1331, A
Sequence 2, Appl
Sequence 2, Appl
Sequence 29, Appl
Sequence 72, Appl
Sequence 8, Appl
Sequence 4, Appl
Sequence 51, Appl
Sequence 174, Appl
Sequence 2, Appl

RESULT 2
US-09-758-269-16

; Sequence 16, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758-269
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO: 16
; LENGTH: 605
; SOFTWARE: Patentin Ver. 2.1
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-758-269-16

Query Match Score 72.4%; DB 10; Length 605;
Best Local Similarity 71.7%; Pred. No. 2.1e-200;
Matches 435; Conservative 63; Mismatches 90; Indels 19; Gaps 6;

Qy 5 TATAAAGSGRLLGCNHTOPPLSSOSSDLSYCSC--SLPMASTYTRKLNVSSAAHTPPAIIH 61
Db 3 TTTSHTATNTW1----KTKLSSPSSKFGFANSISLKNHQRSLNNSLQAPILH 57

Qy 62 FPQOSSN--SPA--IVYKPKAKESN----TKOMNLFQRAAAALDAAEGLVSHFKLHP 112
Db 58 FPKQSSNQTPKNTTSHPKQOENNNSSSSSKWNIVOKAAAMAIDAYEASLTKHLEHP 117

Qy 113 LFKTADPSVQTAGNFVNNEQVRRLNPVYGRLLPDSIKGYVVRNGANPLHEPTGHFFD 172
Db 118 LFKTADPQVLSQGNFAPVPENVCQSLPVTKCIPKVQGVVVRNGANPLFEPTAGHFFD 177

Qy 173 GDDMWHAVKFHGSASAYACRFOTNRFYQERGLGPVKPAIGELIGHTGIAALMFYAR 232
Db 178 GDDMWHAVQFKNGSASTACRFETERLVQEKALGRPFVKPAIGELGHSGTARLMFYAR 237^{*}

Qy 233 AAAGIVDPAHCTGVANAGLVENGRLAMSEDDLPYQWQITPNGDLKTVGREFDFDQLES 292
Db 238 GFGLGLVSKGTVKNGRQLGRPVPAIGELIGHTSTARMLFYARAAAGTVDP 297

Qy 293 TMAHPKVDPESELFAALSYDVSXPKPILKYFRSPDGKSPDVEIQOLDQPTMMHDFATE 352
Db 298 TMAHPKVDPESELFAALSYDVTQKPQKTFYFSEKSNESNEVIEPDPTMMHDFATE 357

Qy 353 NEVVPDQVWSKLPENMRGGSPVWYKVKVAFGLKQYADDSNNKWIADPDCFCFH 412
Db 358 NEVVPDQVVFKMSMRGGSPVYVKKVNSREGTLKYAKGSDLKWEVPDFDCFCFH 417

Qy 413 WNAWEEPDTDEVVVIGSSMTPPDSIFENESDENKSVSLEIRUNLKGTESTRPLISNDQ 472
Db 418 WNAWEEPDTDEVVVIGSSMTPPDSIFENESDENKSVSLEIRUNLKGTESTRPLISNDQ 477

Qy 473 QYNLEAGHVNVRNLGRKTKFAYTALABWPWKVSGFAKVDLTGEVKKHLGDNRGGPL 532
Db 478 QYNLEAGHVNVRNLGRKTTEYALIAEFWPWKVSGFAKVNLTGEVEKFYGDNKYGEPL 537

Qy 533 FLPGE--GGEDEGYLTCEVHDEKTWKSELQVNLNAVEEATVKLPSRVYGFHSTFIG 590
Db 538 FLPRDPNSEEDEGYLIAFVDEKEWKSSELQVNLNAVEEATVKLPSRVYGFHSTFIG 597

Qy 591 ADDLAQK 597
Db 598 ANDLANQ 604

RESULT 3
US-09-758-269-12

; Sequence 12, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758-269
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO: 16
; LENGTH: 605
; SOFTWARE: Patentin Ver. 2.1
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Vigna unguiculata
US-09-758-269-12

Query Match Score 68.8%; DB 10; Length 612;
Best Local Similarity 68.6%; Pred. No. 4e-190;
Matches 423; Conservative 61; Mismatches 98; Indels 35; Gaps 8;

Qy 8 AAVSGRGGNHQTPLPSSOSSDLSYCSCSSLPMASRYTRKLNVSSAAHTPPAIIH 65
Db 3 SSASNNTWNATLPSPPKDPLSTS-SPTNLPL--RKTSSNTICSLQT--LHPPKQY 56

Qy 66 ----SSNSPAIVVKPKAKESN-----KOMNLFQRAAAALDA 100
Db 57 QPTSTSTSTAFTTTPTPKTWTITTPPRETNPLSDTNOFLPKQWNELQKAATALDLY 116

Qy 101 EGFLVSHKPLPKTADPSVQAGNFVNNEQVRRLNPVYGRLLPDSIKGYVVRNGANP 160
Db 117 ETALYSHRKPLPKTADPQVIAGNFVNNEQVRRLNPVYGRLLPDSIKGYVVRNGANP 176

Qy 161 LHEPVGHFFPDGDMHVAYKFEHGSSASYACRFETQNRFLQYERGLGPVKPAIGELGHG 220
Db 177 LYEPVAGHFFPDGDMHVAYKFTNGASAYCRFTERLQEKALGRPFVKPAIGELGHG 235

RESULT 4
 US-09-758-269-2
 ; Sequence 2, Application US/09758269
 ; Patent No. US20020104120A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TUCHI, SATOSHI
 ; APPLICANT: KOBAYASHI, MASATOMO
 ; APPLICANT: SHINOZAKI, KAZUO
 ; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
 ; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
 ; FILE REFERENCE: 3914_3
 ; CURRENT APPLICATION NUMBER: US/09/7758,269
 ; CURRENT APPLICATION NUMBER: US/09/7758,269
 ; PRIORITY DATE: 2001-01-12
 ; PRIORITY APPLICATION NUMBER: JP 2001-003476
 ; PRIORITY APPLICATION NUMBER: JP 2000-010056
 ; PRIORITY FILING DATE: 2000-01-13
 ; PRIORITY FILING DATE: 2000-01-13
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 583
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-758-269-2

Query Match 63.2%; Score 1991; DB 10; Length 583;
 Best Local Similarity 64.5%; Pred. No. 7.2e-174; Mismatches 105; Indels 22; Gaps 6;

Matches 380; Conservative 82; Mismatches 105; Indels 22; Gaps 6;

QY 16 GGNTTQQPPLSSQSSDLSTY -CSSLPMASRVTRKLNSALHTPPQSSNSPATV 73
 QY 11 GGKTTWP--- -QAQTDLGRPIKKPVKCTVQDVTETLTKKRQLFTPTATPP-- 62
 QY 74 VKPRAKESSKOMNLFORAAAALDAAEFLVYSEKLLPMPKTAADPSVYTAGNPAVDEQ 133
 QY 63 ----- -QHPLRLNIFQKAATAIDAAEALISHEQDSLPLKTAADPRYIAGNNSPVPS 116
 QY 134 PVRNLPVQGKLPDSIKGKVYVRNGANPLHEPVTGHFFGDGMHAYKEHGSASYACRF 193
 QY 117 SVRRNLTVESTIPCDGTYIRGANPMPEPTAGHHLFGDGMHAYKLTINGSASYACRF 176
 QY 194 TQTNRREVQEROLGRGVFPKAIGELHGHTGJARMLMFYARAAGIVDPAHGTGVANAGLY 253

RESULT 5
 US-09-758-269-14
 ; Sequence 14, Application US/09758269
 ; Patent No. US20020104120A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TUCHI, SATOSHI
 ; APPLICANT: KOBAYASHI, MASATOMO
 ; APPLICANT: SHINOZAKI, KAZUO
 ; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
 ; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
 ; FILE REFERENCE: 3914_3
 ; CURRENT APPLICATION NUMBER: US/09/7758,269
 ; CURRENT APPLICATION NUMBER: US/09/7758,269
 ; PRIORITY APPLICATION NUMBER: JP 2001-003476
 ; PRIORITY FILING DATE: 2001-01-11
 ; PRIORITY APPLICATION NUMBER: JP 2000-010056
 ; PRIORITY FILING DATE: 2000-01-13
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 604
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-758-269-14

Query Match 61.3%; Score 1930; DB 10; Length 604;
 Best Local Similarity 63.3%; Pred. No. 3e-168; Mismatches 73; Indels 22; Gaps 8;

Matches 377; Conservative 73; Mismatches 122; Indels 22; Gaps 8;

QY 22 PPLSSQSSDLSYCSLPMASRVTRKLNSALHTP-- -PALHFKQSSNSPATV---- 74
 Db 17 PARFARAS---- -NSVRSRSPRAYSVPAECLQAPPVPAVPGHA 71

QY 75 -KPKAKESNTKOMNLFORAAAALDA-AEGFLVS-HEKLHPLPKTADPSVQAGNAPVN 131
 Db 72 AAPKAEKGKKQLNIFQRAAAALDAFEFGFVANLERPHGLISTADAVQAGNAPVG 131

QY 132 EQPVFRNLPLPVGKLPLDSIKGKVYVRNGANPLHEPVTGHFFGDGMHAYKEHGSASYA 190
 Db 132 ERPPVHELPVSPGRIPPFIDGVYARANGANCPDFDPAVGHHEFDGDSMWHALRIRGAESA 191

QY 191 CRFTTNRFVQERQGPVFPKAIGELHGHTGJARMLMFYARAAGIVDPAHGTGVANAG 250
 Db 192 CRFTETARLQEERAIGRPVFPKAIGELHGHSIGJARLALYARAAGLVDPASGTGVANAG 251

QY 251 IYFFNGLRAMSEDILPLPYQVQITPNGLTKVGRPFQDQLESTMIAHPKVDPESGELFAL 310

Query 396 SSNIKWIDAPDCFCFHLWNAAWEETDEVVVIGSOMTPPDSIFNESDENIKSVLSEIRLN 455
 Db 389 ESEMKNFEVPPGENIHTAINAWDEDDGNSVILIAPINTMSIHTLERMD-LHALVEKVKT 447
 Query 456 LKTGSTERRPLTISNEQDQVNLEAGMYNRMNLGRKTFKAYLALAEPWPKVSGFAKVDLTTG 515
 Db 448 LVTGIVYRHPISAR---NLDFAVINPAFLGRCSRYVAAIGDPMPKISGVKLDVSKG 502
 Query 516 E----VKKHLYGDNRYYGGPEPLPGEGLDEDEGYILCFVHDEKTVKSELQIVNAVSLEE- -ATVKLPSRV 566
 Db 503 DRDDCTVARMYGSGCYGGPEFVARDGPNAEEDGYYVTTVHDEVTGESKELVMDAK 562
 Query 567 S--LEYEATVKLPSRVPYGFHGTFIGADDLK 596
 Db 563 SPELEIVAVNRLPRRVYGFHGTFIGADDLK 594

RESULT 8
 US-09-758-269-8
 ; Sequence 8, Application US/09758269
 ; Patent No. US20020104120A1
 ; GENERAL INFORMATION:
 ; APPLICANT: IUCHI, SATOSHI, MASATOMO
 ; APPLICANT: SHINONAKI, KAZUO
 ; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
 ; FILE #: CREENCE: 3914-3
 ; CURRENT APPLICATION NUMBER: US/09/758,269
 ; CURRENT FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: JP 2001-003476
 ; PRIOR FILING DATE: 2000-01-13
 ; PRIOR APPLICATION NUMBER: JP 2001-003476
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: JP 2000-01-0056
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 8
 ; LENGTH: 538
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-758-269-8

Query Match 29.8%; Score 938; DB 10; Length 538;
 Best Local Similarity 37.0%; Pred No. 2.5e-77; Mismatches 105; Indels 48; Gaps 13;
 Matches 207; Conservative 105; Mismatches 199; Indels 48; Gaps 13;

Query 64 KOSSNSPATIVYKPK-AKESENNTKOMMLFQRAAAAALDAEAEGLVSKLPLPKTADPSVQ 122
 Db 4 KLSGDGSSIISVHPRPSKGFFSSKLDDLLERLVVKLM-----Y 47

Query 123 IAGNEAPV-NEOPVRNLPVYKPLDSTIKGYYVRRANGANPLHEPVGHFFDGMHYAVK 181
 Db 48 LSGNFAPIRDETPVYKLPVHGFPLPCCLNGEFVRVSPNPKEDAVAYHWFDGMHGVR 107

Query 182 FEHGASASYACRETQTNRFVQERQLGREVFPAIGEGLHTGJARMLFYARAAGIVDPA 241
 Db 108 IKDGKATYVSRVKTISLKOEEFFGAKFMK-IGDLKGFFGMLMVNQQLRTKLKDNT 166

Query 302 PEGSELEPLSYDVSKPLKYFRESPGCTKSPDVEIQDQPTMMHDFAITENFVYVPPDQQ 361
 Db 227 PVIGEMFTFGYS-HTPPYLTYRVTSKQGTMHDPPVPTISEPIMMHDFAITEYIAFMDFLP 285

Query 342 HGTGVANAGLYYFNGRLAMSEDDLPYQVQTIPNGDLKTVGRFEDQQLESTIMIA:PKVD 301
 Db 167 YGNTANTALVYHKGKLALQADKPYVTKLEDGLQTLGIDYDKRITHSFTA:PKVD 226

Query 362 VFKLPMENTRGGSPV-VDKNKAFTVAFGLDQKTAEDSSMKWIDAPDCFCFHLMNAVEPE 420
 Db 286 MHPRKEMKEKAMISDPTKARFGVLPYRAKDELNMRWFELPNCEFHNNANANG-E 343

Query 421 TDEVVVGSCMTPPD----SIFNESDENIKSVLSEIRLNLTGSTERRPLISNEDQVN 475
 Query 476 LEAVMVNRMLGRKTFKAYLALAEPWPKVSGFAKVD-----LITGEVKKHLY-- 522

Query 477 EDENVVLTICRLENPDIDMVSCKVKEKLENFGNELYEMRNMKTFGSAQOKLSSASA----- 398

Query 478 DEVVVLTICRLENPDIDMVSCKVKEKLENFGNELYEMRNMKTFGSAQOKLSSASA----- 398

Query 479 VDEPRINECYTGKKQKVVYGTLDSSAKVTGLIKEDHAREATGKMLLEVGNKLYIDL 458

Query 523 GDNRYYGSEPLFPLGESEDEGYILCFVHDEKTVKSELQIVNAVSLEE- -ATVKLPSRV 580
 Db 459 GEGRYGEALYVPREAEDDGLLTFVHDETGKSCVTVDAKTSQAEPVAVVEPHRV 518

Query 581 PYGFHGTFIGADDLKQV 599
 Db 519 PYGFHALFVTEEQEQTL 537

RESULT 9
 US-09-758-269-1B
 ; Sequence 18, Application US/09758269
 ; Patent No. US20020104120A1
 ; GENERAL INFORMATION:
 ; APPLICANT: IUCHI, SATOSHI, MASATOMO
 ; APPLICANT: SHINONAKI, KAZUO
 ; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
 ; FILE #: CREENCE: 3914-3
 ; CURRENT APPLICATION NUMBER: US/09/758,269
 ; CURRENT FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: JP 2001-003476
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: JP 2000-01-0056
 ; PRIOR FILING DATE: 2000-01-13
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 18
 ; LENGTH: 538
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-758-269-1B

Query Match 29.7%; Score 937; DB 10; Length 538;
 Best Local Similarity 37.0%; Pred No. 3.1e-77; Mismatches 199; Indels 48; Gaps 13;
 Matches 207; Conservative 105; Mismatches 199; Indels 48; Gaps 13;

Query 64 KOSSNSPATIVYKPK-AKESENNTKOMMLFQRAAAAALDAEAEGLVSKLPLPKTADPSVQ 122
 Db 4 KLSGDGSSIISVHPRPSKGFFSSKLDDLLERLVVKLM-----Y 47

Query 123 IAGNFAPV-NHQPVIRRNLPVYKPLDSTIKGYYVRRANGANPLHEPVGHFFDGMHYAVK 181
 Db 48 LSZNFAPIRDETPVYKLPVHGFPLPCCLNGEFVRVSPNPKEDAVAYHWFDGMHGVR 107

Query 182 FEHGASASYACRETQTNRFVQERQLGREVFPAIGEGLHTGJARMLFYARAAGIVDPA 241
 Db 108 IKDGKATYVSRVKTISLKOEEFFGAKFMK-IGDLKGFFGMLMVNQQLRTKLKDNT 166

Query 242 HGTGVANAGLYYFNGRLAMSEDDLPYQVQTIPNGDLKTVGRFEDQQLESTIMIA:PKVD 301
 Db 167 YGNTANTALVYHKGKLALQADKPYVTKLEDGLQTLGIDYDKRITHSFTA:PKVD 226

Query 302 VYFKLPMENTRGGSPV-YDKNKAFTVAFGLDQKTAEDSSMKWIDAPDCFCFHLMNAVEPE 420
 Db 286 MHPRKEMKEKAMISDPTKARFGVLPYRAKDELNMRWFELPNCEFHNNANANG-E 343

Query 421 TDEVVVGSCMTPPD----SIFNESDENIKSVLSEIRLNLTGSTERRPLISNEDQVN 475
 Query 477 EDENVVLTICRLENPDIDMVSCKVKEKLENFGNELYEMRNMKTFGSAQOKLSSASA----- 398

Query 478 DEVVVLTICRLENPDIDMVSCKVKEKLENFGNELYEMRNMKTFGSAQOKLSSASA----- 398

Db 399 VDFPRINECYTGKQRYVYCTILDSTAKVTGILKEFLHAFETGKRMLEYGGNFKGYIDL 458
 Qy 523 GDNRYGGEPFLPGEDEGGYLICFVHDEKTWSLQIVNAVEVE-ATVKLPSRV 580
 Db 459 GDNRYGGEPFLPGEDEGGYLICFVHDEKTWSLQIVNAVEVE-ATVKLPSRV 580
 Qy 581 PYGFHGTFIGADDLAKQVW 599
 Db 519 PYGFHALFVTEQLQEQT 537

RESULT 11
 US-08-976-063C-22
 Sequence 22, Application US/08976063C
 Publication No. US2002018269 A1
 GENERAL INFORMATION:
 APPLICANT: Alexander Steinbuchel; Horst Briefert; Jurgen Rabenhorst
 TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF
 CONIFERYL ALCOHOL, CONIFERYL ALDEHYDE, FERULIC ACID, VANILLI
 TITLE OF INVENTION: ACID AND THEIR USE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE
 STREET: 660 White Plains Road
 CITY: Tarrytown
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10591-5144
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
 COMPUTER: HP VECTRA
 OPERATING SYSTEM: DOS
 SOFTWARE: WordPerfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,063C
 FILING DATE: 21-NOV-1997
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 196 49 655.1 (Germany)
 ATTORNEY/AGENT INFORMATION:
 NAME: Kurt G. Briscoe
 REGISTRATION NUMBER: 33,141
 REFERENCE/DOCKET NUMBER: Bayer 9998-CAO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 332-1700
 TELEFAX: (914) 332-1844
 TELEX:
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 505 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-063C-22

Query Match 29.7% Score 936; DB 10; Length 538;
 Best Local Similarity 37.0%; Pred. No. 3.9e-77; Gaps 13;
 Matches 207; Conservative 105; Mismatches 199; Indels 48; Gaps 13;

Qy 64 KQSSNSPAIVVKPK-AKESNTKOMNLORAAGAALDAAEGLVLSHEKLHPLEKTADPSVQ 122
 Db 4 KLSGDSTTISVPSRPSKGFFSKLDSLLERYVKL-----Y 47

Qy 123 IAGNFAP-NEQVPPVRLPVGVLPDTSKGVVVRNGANPLHEDDGMYHAVK 181
 Db 48 LSGNEAIRTDEEPPVLPVHOFPLPECUNGEVRYGNPKEDAVAYHWEFGMWHGVR 107

Qy 182 FPHGSASACRFTQTNRNFRVQERQLGREPFKAIIGELGHHTGIAIRMLFYARAAGIVDPA 241
 Db 108 IEDKGKAVYVSRVTKTSRIKQEFFGARKFMK TGDJKFGFLMVNVQLRKLLDNT 166

Qy 242 HGTGVAAGLYENGRILLAMSDDLPTFOVITPNGLKTVGRFEDPGOLESTMIAHKVD 301
 Db 167 YGNTANTALVYHGHKLIAQADPKVVKIVLQDGLQTGLTIDYDRLTHSFTAHKVD 226

Qy 302 PESGEELALSYDVSKPKYIKEYRFSPOTSVDVIEQDQPMHMDFATIENVVFDQO 361
 Db 227 PTGEMETFGYS-HTPPYLTXVYLSKOGIMHDPPVTTISEPMHMDFAITEYTAIFNDLP 285

Qy 362 VPKLPMEIRGGSPVV-YDKNVAAREGLDKYAEDSSNIKWLDAFDGCFHWNAWBEPB 420
 Db 286 MHPRKPKVNEKKEKMMIYSDPPTKARFKGVLPAKDELWPELFWHANANE-E 343

Qy 421 TDEVVIGSCMTPPD----SFNESENDSLNLKVSETRLNKIGESTRRPITSNEDQVN 475
 Db 344 EDEVVLTICRLENFDLDVSGKYREKLFENGFNEYMRNMIGSASOKLASA----- 398

Qy 476 LEAQMVNRNMLGKRTKFAYLAALAEPPDKVSGFAKVD-----LITGVKKHLY- 522
 Db 399 VDFPRINECYTGKQRYVYGTILDSTAKVTGILKEFLHAFETGKRMLEYGGNFKGYIDL 458

Qy 523 GDNRYGGEPFLPGEDEGGYLICFVHDEKTWSLQIVNAVEVE-ATVKLPSRV 580
 Db 459 GDNRYGGEPFLPGEDEGGYLICFVHDEKTWSLQIVNAVEVE-ATVKLPSRV 580
 Qy 581 PYGFHGTFIGADDLAKQVW 599
 Db 519 PYGFHALFVTEQLQEQT 537

RESULT 11
 US-08-976-063C-22
 Sequence 22, Application US/08976063C
 Publication No. US2002018269 A1
 GENERAL INFORMATION:
 APPLICANT: Alexander Steinbuchel; Horst Briefert; Jurgen Rabenhorst
 TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF
 CONIFERYL ALCOHOL, CONIFERYL ALDEHYDE, FERULIC ACID, VANILLI
 TITLE OF INVENTION: ACID AND THEIR USE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE
 STREET: 660 White Plains Road
 CITY: Tarrytown
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10591-5144
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
 COMPUTER: HP VECTRA
 OPERATING SYSTEM: DOS
 SOFTWARE: WordPerfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,063C
 FILING DATE: 21-NOV-1997
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 196 49 655.1 (Germany)
 ATTORNEY/AGENT INFORMATION:
 NAME: Kurt G. Briscoe
 REGISTRATION NUMBER: 33,141
 REFERENCE/DOCKET NUMBER: Bayer 9998-CAO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 332-1700
 TELEFAX: (914) 332-1844
 TELEX:
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 505 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-063C-22

Query Match 13.1% Score 411.5; DB 1; Length 505;
 Best Local Similarity 25.8%; Pred. No. 4.6e-29;
 Matches 137; Conservative 79; Mismatches 192; Indels 123; Gaps 19;

Qy 118 DPSVQIAGNFAPNEQVPRRNLPVVKLPDTSKGVYVRNGANPLHEDVGMYHFRGDMV 177
 Db 8 DP-QLVQTLPLPRIEADLFDLEVGEPKSLSINGTFRTNPQVTPQKFHTFIDGDM 65

Qy 178 HAVKFEHGASYASACRETQTNRFVQERQGRPVFPKATEGLHHTGIALRMLFYARAAGI 237
 Db 66 SAPFHFEDHVDI-SRWTWYKTAERTAERLKSKJEGMTRNPYDDTSVKG----- 114

Qy 238 VDPAHGTGIVYANAGLYENGRILLAMSDDLPTFOVITPNGLKTVGRFEDPGQLES-TMIA 296
 Db 115 -----DRVANTSIISIHKGKVLAKEGLP-YELDR-TLETRGHDYDGSYTSQHTA 166

Qy 297 HKVDFPESGELFALSVDVSKPKYFVPSGOTKSBDV-----EIQLQDQ- 342
 Db 167 HKYDPEGDL-----FFGSAAKEATFDMAVYIVDKHKGKVTHETWFEEQPY 213

Tue May 20 09:35:02 2003

us-09-758-269-6.rapb

Page 9

Db	16244	KAPEVTTVKDSMIVWWERPASDGSEITLGYYLEKRDKEGIRWTCHKRLLIGELRLRV-T	16302
QY	459	GESTRRPIISNEDQQVNILEAGMVNRMLGKTKFAYLAALEPWPWPKYSGFKR	-509
Db	16303	G-----LIEENHDYEFVSA----ENAAG-----LSEPSPPP-SAYQACDPYKPG	16342
QY	510	-----VDLTGEV---KKHLYGDNRGGBPLFLPGESEEDGYIL-CFVH-DEK^	555
Db	16343	PPNNRKVIDTRSSVPLSWSKPIY-----DGCGCTGFTIVEXCDVNGETW	16388
QY	556	WKSELQIVNAYSLEVEATVKLPSRYPYGF	584
Db	16389	MCTPPPTGINKTNEVE--KILLEKHEYNF	16414

Search completed: May 19, 2003, 15:58:18
Job time : 77 secs

Gencore version 5.1.4_p5_4578
 Copyright (c) 1993 - 2003 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: May 19, 2003, 15:46:15 ; Search time 21 Seconds
 (without alignments)
 2742.120 Million cell updates/s
 title: US-09-758-269-6
 perfect score: 3150
 sequence: 1 MASFTAAVSGRNLGGNHT.....VBYGFHGTFIGADDLAKQVV 599

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283242 seqs, 9613422 residues
Total number of hits matching chosen parameters:

chosen parishes: St. Saviour's, St. George's, St. Paul's, St. John's, St. Peter's, St. Michael's, St. James' and St. Anne's.

minimum DB sea length: 9

Maximum DB seq length: 200000000

post-processing: Minimum Match 0%

Maximum Match 100%
Match 100%
Match 100%

Listing first 45 summaries

database : PIR 73:*

```
3 : Pir3:*
```

4: pir4:#

and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	2289	72.7		604	2	T51936		probabilistic
2	2280.5	72.4		605	2	T07123		nine-nine
3	2170	68.9		657	2	E96812		proto-
4	2044.5	64.9		589	2	A86425		proto-
5	1991	63.2		583	2	T04531		nine-nine
6	1930	61.3		604	2	T04351		viviparous
7	968	30.7		595	2	T04438		hypothetical
8	938	29.8		538	2	T49193		neoxarid
9	937	29.7		538	2	T51734		neoxarid
10	745.5	23.7		475	2	AG1944		hypothetical
11	745	23.7		446	2	T17019		hypothetical
12	528.5	16.8		483	2	E87345		conserved
13	466	14.8		501	2	A70534		hypothetical
14	454.5	14.4		502	2	A70582		hypothetical
15	394	12.5		490	2	S76169		hypothetical
16	355.5	11.3		497	2	AB2341		lignocellulose
17	352	11.1		484	2	JN0595		lignocellulose
18	317	10.1		480	2	S76206		hypothetical
19	303.5	9.6		489	2	JC4324		lignocellulose
20	253	8.0		616	2	T10688		hypothetical
21	244	7.7		618	2	C84985		hypothetical
22	242	7.7		472	2	AG2417		hypothetical
23	238.5	7.6		556	2	F88115		proto-
24	231	7.3		483	2	D87290		conserved
25	215	6.8		533	2	A47143		retinol
26	126	4.0		790	2	T34293		hypothetical
27	116.5	3.7		1832	2	T31113		mucin
28	114	3.6		870	2	S74291		hypothetical

ATTACHMENT

RESULT 1
 T51936
 probable 9-cis-epoxycarotenoid dioxygenase [imported] - potato
 C;Species: Solanum tuberosum (potato)
 C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C;Accession: T51936
 R;Burdidge, A.; Taylor, I.B.; Thompson, A.
 submitted to the EMBL Data Library, March 2000
 A;Description: Potato Putative 9-cis-epoxycarotenoid dioxygenase 1 cDNA.
 A;Reference number: 225874
 A;Accession: T51936
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: mRNA
 A;Residues: 1-604 <BUR>
 A;Cross references: EMBL:AJ276244; PIDN:CAB76920.1
 C;Genetics:

Db	418 NAWEEPETDEIVVIGSCMTPPDSTFNEDGEGLKSVLSEIRLNKTGKSTRKAIKENPDEQ	477	Db	538 FLPRDPNSKEEDGYTIAFVHDEKEWKSEELQIVNAMSLKLEATVKLPPSRVPYGFHGTEIN	597
Qy	474 VNLEAGMVRNMMLGRKTPEAYALAAEPKPKVSFAKVDLTGEGVKKHLYGDNNGEGLF	533	Qy	591 ADDIAKO	597
Db	478 VNLEAGMVRNMMLGRKTPEAYALAAEPKPKVSFAKVDLTGEGVKKHLYGDNNGEGLF	533	Db	598 ANDIANQ	604
Qy	534 LPGE--GGEDEGYILCPVHDDETWKSPELQIVNAVSLEAVTKLPSPVYGFHGTFGA	591			
Db	538 LPDPNSKEEDGYTIAFVHDEBEWTSELDQIVNAVTMILKLEAVTKLPSPVYGFHGTFNA	597			
	RESULT 3				
	E96812				
	protein F3F9_10 [imported] - Arabidopsis thaliana				
	C;Species: Arabidopsis thaliana (mouse-ear Cress)				
	C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001				
	C;Accession: E96812				
	R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alan				
	Chin, C.W.; Chaudhuri, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Dewar,				
	anssen, N.F.; Hughes, B.; Huizar, L.				
	Nature 408, 816-820, 2000				
	A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,				
	C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia				
	Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo				
	ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venier, J.C.; Davis, R.W.				
	A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.				
	A;Reference number: A86141; PMID:21016719; PMID:11130712				
	A;Accession: E96812				
	A;Status: Preliminary				
	A;Molecule type: DNA				
	A;Residues: 1-657 <STOP>				
	A;Cross-references: GB:AE005173; NID:98052533; PIDN:AAF71797_1; GSPDB:GN00141				
	A;Gene: F3F9_10				
	A;Map Position: 1				
	Query Match				
	Best Local Similarity				
	Matches 435; Consistency 71.7%; Pred. No. 6e-165;				
	Matches 419; Mismatches 90; Indels 19; Gaps 6;				
	Query Match				
	Best Local Similarity				
	Matches 419; Conservative 80; Mismatches 80; Indels 33; Gaps 9;				
Qy	5 TATAAAGSRRWLGGNHTOPPLSSOSSDLSYC	61	Qy	1 MASFTATAAVGRWL-----GGNHHTQPLSSQSSDLSYCSSLP-----MASRTV 45	
Db	3 TTTSHTATNTWI-----KTKLSPMSKEFGAANSISLKNQHNRQLNNSLQAPPILH	57	Db	51 MASTTILPSTSTQFLDRTFTSSSSSRPKLQS-----LFSSTLNKKLVPCYSSVN 105	
Qy	62 EPKQSSN-SPA- -IVYKPKAKESN----TKQMNIFQRAAAALDAEGFLVSHIELKHP	112	Qy	46 RKLNVSSALHPT----PPAHLPKQSSNSPAPAVVKPKAKESNTKQNNLFQRAAAALDAE 101	
Db	58 FPQKSSNYYQTPKNNNTSHPKQENNNSSSSTSNNVOKAAMADAVESALTAKIELEHP	117	Db	106 KKSYSVSSLQSPPTFKPSW-----KKLCNDVNLNP-TKTNONPK-LNPVQTAAMVLDATE 160	
Qy	113 LPKTADEPSVQLAGNEAPVNQPVRLNPVYKGPKLPSIKGVYVRANGANPLHEPTGHFFD	172	Qy	102 GFLPYTHGHFEDGDMVHAYKFEHGSASACRFTOTNRFVQEROLGRPVPEPKAIGELRH 220	
Db	118 LPKTADEPSVQLAGNEAPVNQPVRLNPVYKGPKLPSIKGVYVRANGANPLHEPTGHFFD	177	Db	161 NAMISHERRHRPHPKTADPAVQIANNFPEPKVPHNPFVGTGTYPECLQGVYVRNGANP 220	
Qy	173 GDGMVHAYKFEHGSASACRFTOTNRFVQEROLGRPVPEPKAIGELRH 232		Qy	161 LHEPYTHGHFEDGDMVHAYKFEHGSASACRFTOTNRFVQEROLGRPVPEPKAIGELRH 220	
Db	178 GDGMVHAYQPKNGNSASYACRFETETRLVQEKAALGRGPVPEPKAIGELGHSSIAIRMLFYAR	237	Db	221 LHKPVSGHII_FGDGMVHAYRFDNSVSVACRFETPNRLVQERECGRPFPKAIGELRH 280	
Qy	233 AAAGTVAHGTTGYNAGAGLYFNGRLLAMSDFLQYQVQITPNSDQKTYGREDDEGQLES	292	Qy	221 TGIALMLMFYARAANGIVDAHGTVGANAGLVYFNGRLAMSDDLPYQVQTIPNGDIFT 280	
Db	238 GLFLGLDHSGTGTVGANAGLVYFNGNLANSSEDFLPYHVAVTPGDLKTMGRFDGQQLKS	297	Db	281 LGIAKLMFNTGFLGLDPTGGLVFNGHLLAMSDDLPYHVKVTOQTGDLT 340	
Qy	293 TMIAPKVDPESGEFALSVDYVSKPVLKYFRESPDGTQSPDVEQLDQPTMMDFATE	352	Qy	281 VGRFDEFQGEQESTMAHPYDPESELSYDVSXKPLPEMIRGSSPVYDKNLYKARFGJLKDQYADESSNK 400	
Db	298 TMIAPKLDEPSGEFALSVDYVSKPVLKYFRESPDGTQSPDVEQLDQPTMMDFATE	357	Db	341 SGRYTDFDGQIKSTMIAHPYDPESELSYDVSXKPLPEMIRGSSPVYDKNLYKARFGJLKDQYADESSNK 400	
Qy	353 NFVVYPDQQVYKLPEMIROSPPVYDKNLYAEDDSNIKWTIDPDCFCFHL 412		Qy	341 QPTMMHDFATTENFVVFVPPDQQVYKLPPEMIRGSSPVYDKNLYKARFGJLKDQYADESSNK 400	
Db	358 NFVVYPDQQVYKLPEMIROSPPVYDKNLYAEDDSNIKWTIDPDCFCFHL 417		Db	401 QPTMMHDFATTENFVVFVPPDQQVYKLPPEMIRGSSPVYDKNLYKARFGJLKDQYADESSNK 400	
Qy	413 WNAWEEPETDEVVIGSCMTPPDSTFNEDENLKSVLSEIRLNKTGESTRRLISNEDQ	472	Qy	401 WIADPDCFFFHLLNAWEETBTDEVVIGSCMTPPDSTFNEDENLKSVLSEIRLNKTG 460	
Db	418 WNAWEEPETDEVVIGSCMTPPDSTFNEDENLKSVLSEIRLNKTGSTRKSTIENDE	477	Db	461 WIEVPDCFFFHLLNAWEETBTDEVVIGSCMTPPDSTFNEDENLKSVLSEIRLNKTG 520	
Qy	473 VNLEAGMVRNMMLGRKTPEAYALAAEPKPKVSFAKVDLTGEGVKKHLYGDNNGEPL	532	Qy	461 STRRPISTSEDOQVNLEAGMVNNMLGRKTPEAYALAEPPKVSFAKVDLTGEGVKKH 520	
Db	478 VNLEAGMVRNMMLGRKTPEAYALAAEPKPKVSFAKVDLTGEGVKKHLYGDNNGEPL	537	Db	521 STRRRVIS--ECVNLAEQMVNRNLGRKTRYAYLATTEPWPKVSFAKVLSTGEIRKX 577	
Qy	533 FLPGE--GGEDEGYILCPVHDDETWKSPELQIVNAVSLEAVTKLPSPVYGFHGTFIG	591			

Qy	521	LYGDNRYYGGPPLFLPGEQQBEGEYTLCFVHDEKTWKSEIOTQYNASLEVATVRLPSRV	580		Qy	565	AVSLEVEATVRLPSRVPYGFHGTFIGADDLAKO	597	
Db	578	TYGKCYKGPPLEPLSGDFEDGYTAVFVHDEEKVKSEIQLINAVNMKLEATVRLPSRV	637		Db	556	AVTLEATEVRLPSRVPYGFHGTFVNSADMLNQ	588	
Qy	581	PYGFHGTFIGADDLAKO	599						
Db	638	PYGFHGTFISKEDELSKQAL	656						
		RESULT 5							
	A86425	nine-cis-epoxycarotenoid dioxygenase homolog F28J12.10 - Arabidopsis thaliana							
		N;Alternative names: hypothetical protein T9A21_200							
		C;Species: Arabidopsis thaliana (mouse-ear cress)							
		C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001							
		C;Accession: A86425							
		R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kau, S.; White, O.; Alonso, C.;A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maiti, R.; Marijalli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.							
		Nature 408: 816-820, 2000							
		A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;Rizzini, M.; Rooney, T.; Rowley, D.; Sakano, H.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;Hager, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venier, J.C.; Davis, R.W.							
		A;Title: Sequence and analysis of chromosome 1 of the Plant Arabidopsis.							
		A;Reference number: A86414; PMID:1130712							
		A;Accession: A86425							
		A;Status: preliminary							
		A;Residue: 1-589 <STOP>							
		A;Cross-references: GB:AE005172; NID:911094779; PIDN:AAG29711.1; GSPDB:GN00141							
		C;Genetics:							
		A;Map position: 1							
		Query Match 64.9%; Score 2044.5; DB 2; Length 589;							
		Best Local Similarity 68.1%; Pred. No. 4.7e-147;							
		Matches 390; Conservative 69; Mismatches 101; Indels 13; Gaps 4;							
Qy	27	SQSSDLSYCSSLPMASRYTRKLASSALHTPPALHFKOSNSNPARYVVKPARESNTKOM	86		Qy	16	GGNHQQPLSSQSOSDSLW-CSSLPMASRYTRKLASSALHTPPALHFKOSNSNPAYV	73	
Db	27	SPSSSVSETNTKPR---BRKLSANSVSDPTNLINFPNYPSPNPII----PEKDTSRW	76		Db	11	GG1ITWP---QAQ1DGLGRPKRPIKQPKVTKCTVQIDVTE-LTKKRQLQFTPTTTATP	62	
Qy	87	NLFQRAAAAALDAEAEGLPVSHKEKLHPKTADEPSVQTAGNFAPVNEQPYVRNLPVVGKLP	146		Qy	74	VKPKAKESNTKOMNLFQRAAAAALDAEAEGLPVSHKEKLHPKTADEPSVQTAGNFAPVNEQ	133	
Db	77	NPLQRAAAALDAEAEGLPVSHKEKLHPKTADEPSVQTAGNFAPVNEQPYVRNLPVVGKLP	146		Db	63	-----OHNPLRNLNIFQRAAAAALDAEAEGLPVSHKEKLHPKTADEPSVQTAGNFAPVNEQ	116	
Qy	147	DSTKGVYVNRGANPLHEPYTGHHFFDGDGMVHAKPFENGASAYACRFQTINRQEVQEROLG	206		Qy	134	PVRLNLPSVQKLPDPSIKGIVYVRNGANPLHEPYTGHHFFDGDGMVHAKPFENGASAYACRF	193	
Db	137	DCIDGVYVNRGANPLHEPYTGHHFFDGDGMVHAKPFENGASAYACRFQTINRQEVQEROLG	196		Db	117	SVRNNTLVETGTTIPCIDGVYTRNGANDMEPTAGHLLFDGDGMVHAKFTTNGASAYACRF	176	
Qy	207	RPVFPKAEGELHGTGIGRLMLFYARAAGIVDPAHGTGVANGLIVENGRLJAMESDDL	266		Qy	194	TQTNRFVQERQLGRPVPKAEGELHGTGIGRLMLFYARAAGIVDPAHGTGVANGLIVY	253	
Db	197	SPIFPKAEGELHGTGIGRLMLFYARAAGIVDPAHGTGVANGLIVENGRLJAMESDDL	256		Db	177	TKTERLQEKRLGPVFPKAEGELHGTGIGRLMLFYARAAGIVDPAHGTGVANGLIVY	236	
Qy	267	PYQVQITPNGLDKTVGRDFDQGLESTMTAHPKVDPESGELFALSVDKPYLKYFRFS	326		Qy	254	FNGRLAMSDDLDLPIQVQZTPNQSDLKLTVQGKDFGQLESTMTAHPKVDPESGELFALSVD	313	
Db	257	PYQVRTONGLETFIGRDDEGQOLSSAMAHAPKIDPVTKELPLALSVDKPYLKYFRFS	316		Db	237	FNNRLAMSDDLDLPIQKLTQGTDQVTPMHDPTVKEHLASYD	296	
Qy	327	PDGTSKSPDVIEQDQPTMMHDFAITENFVYVDPDQVYVKEPLPEMIRGSPVYVDPDNKLYARF	386		Qy	314	VVKSPYDKNKVAREFGILDKYAEDDSNTKWDADPCFCFHLMNAWEPEPTDEVVYVIGSOMTP	433	
Db	317	PEGEKSPVPEVPIAPSTMHMDFAITENFVYVDPDQVYVKEPLPEMIRGSPVYVDPDNKLYARF	376		Db	337	SPVFDGKVKSRGLIMPQDATAQSIQIWNNSPTECHLWNNAWEPEPTDEVVYVIGSOMTP	416	
Qy	387	GILDKYAEDDSNTKWDADPCFCFHLMNAWEPEPTDEVVYVDPDNKLYARF	446		Qy	434	PDSIENESDENLKVSLESLRSLVNLKGTGESTRPLISNDQVNLEAGMVNRNLGRKTFA	493	
Db	377	GILPRNKADESNWVWESPETPCFHLMNAWEPEPTDEVVYVDPDNKLYARF	436		Db	417	ADSIFNERDESLRSVLESLRSLVNLKGTGESTRPLISNDQVNLEAGMVNRNLGRKTFA	474	
Qy	447	SVLSEIRNLKGTGESTRPLISNDQVNLEAGMVNRNLGRKTFA	506		Qy	494	YLALAEWPKVSGFAKVDLTTGEVKHLGYDNRGGPLFLPGEGL--EDEGYILCFL	549	
Db	437	SVLSEIRNLKGTGESTRPLISNDQVNLEAGMVNRNLGRKTFA	495		Db	475	FLAIAYWPKVSGFAKVDLCTGEMKKVYGGKGGPEFLPGEGL--EDEGYILCFL	534	
Qy	507	FAKVDLITGEVKHLGYDNRGGPLFLPGEGL--EDEGYILCFL	564						
Db	496	FAKVDLSTGEVKHLGYDNRGGPLFLPGEGL--EDEGYILCFL	555						

QY 550 VHDEKTKWSELQLTINYNAVSLEVEATVKLPSPRVPGFHTGFTGADDLAKQV 598
 Db 535 VHDETKTSELQLTINYNAVKIEATIKLFSRVPFGFHTGVDSNELVQD 583

RESULT 6

T04351 6 viviparous-14 protein - maize
 C;Species: Zea mays (maize)
 C;Accession: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
 R;Tan, B.C.; McCarty, D.R.
 Submitted to the ENBL Data Library, March 1997
 A;Reference number: Z15304
 A;Accession: T04351
 A;Status: preliminary; translated from GB/ENBL/DBJU
 A;Molecule type: mRNA
 A;Residues: 1-604 <TN>
 A;Cross-references: EMBL:U95053; NID:92232016; PIDN:AAB62181.1; PID:92232017
 A;Experimental source: strain W22
 C;Genetics:
 A;Gene: vp14
 A;Map position: 1
 C;Function:
 A;Description: catalyzes oxidative cleavage of 9-cis-epoxy carotenoids to yield xanthoxin

Query Match 30.7%; Score 968; DB 2; Length 595;
 Best Local Similarity 36.7%; Pred. No. 2.5e-65;
 Matches 210; Conservative 106; Mismatches 220; Indels 36; Gaps 13;

QY 42 SRVTRKLNVSSALHTPPALHPFKOSSNSPAIVVKPKAKESNTKOMNLFORAAAALDAE 101
 Db 42 SPVNPSNDNRRPKTLH--NRTNTHFLVSSPKLKPEMTLATAFL---TTVEDWIN 94

QY 102 QFLVSHEKLHPPLPKTADPSVQYTAGNFAPVNEQPYRRNLPV-GRLPDTSKGIVYVRNGANP 160
 Db 95 TFIDP---PSRPSVDPHVLSDNFPVLDPLDELPPTDDEIIHGTLPLSTNGAVIRNGNP 149

Query Match 30.7%; Score 968; DB 2; Length 595;
 Best Local Similarity 36.7%; Pred. No. 2.4e-138;
 Matches 210; Conservative 106; Mismatches 220; Indels 36; Gaps 13;

QY 161 LHEPVTVGHFFFDGDMVHAYKFEGHSASACRFETQTNRFVQEROLGRPFVKAGELGH 220
 Db 150 QFLPRGPYHLFDGDMHLAIIKHNGKALCWSRIVKTYKINEROTQGAPVMPNVESGFNGV 209

QY 17 PARSHARAS---NSVRFSPRAYSSVPAECQAFPKPVAIDLPPRKPAIAVCPHA 71
 Db 17 PARSARAS---NSVRFSPRAYSSVPAECQAFPKPVAIDLPPRKPAIAVCPHA 71

QY 75 -KPKAKESNTKOMNLFORAAAALDA-AEGFLYS-HEKLHPPLKPTADPSVQZAGNFAPVN 131
 Db 72 AAPSKRAEGKKQNLNLFORAAAALDAEEFGVANVLERPHGLSTADAEGVZAGNFAPVG 131

QY 132 EOPYRNLIVQKLPDSDTRGVYVNGNHLPEPQVHIFGKDFDMVHAKFENGSA-SYA 190
 Db 132 ERPVHHELPVSGRPPIPFDGVYANGANCFDPVAGHILFDGGMVHALRNGAAESYA 191

QY 191 CRFTQTNRFVQEROLGRGPVPEPKAIGEYIAGHTGIARMLFYARAAGITVDPAHTGVANAG 250
 Db 192 CRFTETARIQERATGRIPEPKAIGEYIAGHTGIARLAUFLYARAACGLYDPSACTGVAG 251

QY 251 LYVNGRLIAMSSEDDLPYQVITPNGLDKTGVRFEDSOLESTMIAHKVDPESGEFL 310
 Db 252 LYVNGRLIAMSSEDDLPYHVRVADGDLTETVGRDFDGLGCAIMIAHKLDPATGELHAL 311

QY 311 SYDVSKPKLYKPFSPDQTKSPVIEQLDQPMNMHDILETTEVQVQVFEKLPMI 370
 Db 312 SYDVTKRPVLYKPFSPDQTKSDDVETPEQPMIHDAITENVVYDHQVVFKLQEMI 371

QY 371 RGGSPPVYTDOKNKVAFGLDKYABDSSNIKWIADPDCFCHLWNAAWEPEPETDVVVGSC 430
 Db 372 RGGSPPVYTDOKNKVAFGLDKYABDSSNIKWIADPDCFCHLWNAAWEPEPETDVVVGSC 431

QY 431 MTPDSIENESDENIKSYLSEIRNLKTGESTRPLISNEDQQVNLEAGMVNRNMLGKRT 490
 Db 432 MTPDSIENESDENIKSYLSEIRNLKTGESTRPLISNEDQQVNLEAGMVNRNMLGKRT 490

QY 491 KEAYLALARPKPKSGFAKVDLTGEVKHLYGDNRYCGEPFLPGESE---EDGY 545
 Db 491 RYAYLAVAPPKPKSGFAKEDLSIGELTKFEYGGFRGSGEPVMPDAAHPRGEDGY 550

QY 546 ILCPVHDEKTWKSELQIVNAVSLEVEATVKLPSPAVYGFHTGFIGADDIAQ 597
 Db 551 VLTYPVHDERAGTSELVLYNADIRLTVQPSRVPVGFHGFITGQLEAQ 602

RESULT 8

T49193 neoxanthin cleavage enzyme ncl - Arabidopsis thaliana
 N;Alternate name: Protein MAA21_150
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Accession: T49193
 R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Newes, H.W.; Rudd, S.; Lemcke submitted to the Protein Sequence Database, April 11 2000
 A;Reference number: Z25018
 A;Accession: T49193

Db	22 LEGNFAVPHBEEITTDLKVGELPELSCMFVRNGPNPWTPIGQYHWFDDGMLHGVRJ	81	Db	312 R-----NLDAVINPAYLGKRNKTVYAAEDPMRPISSGVKLDSNVRECIASRMFG	366
Qy	183 EHGSAKYACRFQTQHFRVQEROLGPVPPFKAGLHLGHGFIARMLFYARAAGLVD-	239	Qy	524 DNRYGCEPLPLPGE---GGEEDEGSYILCFVHDEKTWKSELSQIVNAVS--LEVEATVKLP	577
Db	82 SNGKTYTNRYYVTRWOLEHAEQAIW-----TIGMEEPPQ	117	Db	367 PGCYGGEPFVAREPENPEADEDNLFSVYHDERAGESFLYMDAKSPOLDIVAAVRMP	426
Qy	240 ---PAHTGYANAGLIVYFNGRLANSSEDDLPYQVOITIPNDLKVGREDDEGOESTRMLA	296	Qy	578 SRVPYGFHGTEFIGADDL	594
Db	118 TELPVNTG -NTLWHAGQLLWEGGAYQVPP--DLASIGEYIYNQSSAFTA	172	Db	427 RRVPYGFHGTEFIGADDL	443
Qy	297 HKVDPESGEFALSAFDVSKPYLKVYFRESPDGTGKSPDVBIEQDLOOPTMMHDAITENFVV	356			
Db	173 HKVDPESGEFALSAFDVSKPYLKVYFRESPDGTGKSPDVBIEQDLOOPTMMHDAITENFVV	231			
Qy	357 VPDQVWVKLEMIRGGSPVYDKVKAVERGILDKYAFDSSNIKIKDADPCCFHUNAW	416			
Db	232 EMDLPLTFESVBRMRQGPMLFESDRPSREGILPRHG-DNSQIROFEAPSCYVFITLNAY	290			
Qy	417 REPEDEVVVIG----SCMTPPSLIFNEDENIK-SVLSSEIRNLKTSGETSTRPLIS	469			
Db	291 E -DKEVYVLFACPMRSTVLAASPQ---QTDPEADIPRHRWRPLKGKVKHEEML---	342			
Qy	470 EDQQVNLEAGMVNRNMLGRKTKFAYTA-LAE-PWPKVSGFFAKVDTTGEYKHLVGDNRV	527			
Db	343 -DDVASESEPRINENELLGQPQYGTTSRILAKGSIPLFEGLIKYDLSNAKSONYEYGQR	400			
Qy	528 GCEPLPLPGEGE-EDEGYILCFVHDEKTWKSELSQIVNAVSLEEV-A TYKLPSPRVYGF	584			
Db	401 GSEAVYVPRPQATEDDGWLTIVYVDTGEESSELVINAQDINSEPIARYVLLPQRVYGF	460			
Qy	585 HGTEFIGADDL	594			
Db	461 HGIWVTEQL	470			
RESULT 11					
Hypothetical protein - apple tree					
C.Species: Malus domestica (apple tree)					
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000					
C.Accession: T17019					
R.Watillon, B.; Kettmann, R.; Arredouani, A.; Hecquet, J.; Burny, A.					
Plant Mol. Biol. 36, 909-915, 1998					
A.Title: Apple messenger RNAs related to bacterial lignostilbene dioxygenase and plant s					
A.Reference number: Z:88655; MUID:98179104; PMID:9520281					
A.Accession: T17019					
A.Molecule type: mRNA					
A.Experimental source: cv. McIntosh, strain Wijcik					
A.Cross references: ENBL:Z93765; NID:92924324; PID:CA07784-1; PID:g2924325					
A.Status: preliminary; translated from GB/EMBL/DDJB					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					
Best Local Similarity 39.8%; Conservative 150; Mismatches 76; Indels 9;					
Query Match Score 23.7%; Pred. No. 1.3e-48;					

RESULT 13
 A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Species: Mycobacterium tuberculosis
 A:Accession: A70534
 C:Accession: A70534
 C:Date: 17-Jul-1998 #text_change 22-Oct-1999
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Gentles, S.; Hamlin, N.; Holroyd, S.; Ratandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squires, S.; Nature 393, 537-544, 1998
 A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70534
 A:Accession: A70534
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Cross references: GB:295972; GB:AL123456; NID:q3261790; PIDN:CA89380.1; PID:e319166;
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv0654

Query Match	14.88;	Score 466;	DB 2;	Length 501;
Best Local Similarity	27.98;	Pred. No. 2.e-27;		
Matches	148;	Conservative	76;	Mismatches 195; Indels 112; Gaps 18;

QY 123 IAGNFAPVNEQQPVRNLPPVGKLPDSTIKGVYVRNGANPLHE --PVTHHHFGRDGMVHAY 180
 Db 14 LEGFLAFYSTEVTATDLPVTRGPEHLDGRLYLRNGPNPVAEVDAT-YHWFPGDAWVHG 72
 QY 181 KFEHGSASYACRFQTQNRFYQERQLGRPFKAIGE ---LHIGHTGIARLMLEYARA 235
 Db 73 ALRDGKA -----RWRYNRWRVTPAVACALGEPISARHPRTGIE-----112
 QY 236 GITVPAHTGVANAGLYFNMGRLAMSEDDLPYQVQITPNGDLKTYGRFDGQLESTM 295
 Db 113 -----GEPNTNVVTHAGRILALVEAGV --VNYLETIDELDTVSPCDDEGTGLHSGYT 160
 QY 296 AHPKVDPESGELEFALSVDVSKPVLKFYFRSPDGTKSPDVEIQLDQPTMMHDFAITENFV 355
 Db 161 AHFQRDPHTGEHLAVSYSFARGHRVQZSVIGTDGHARTVDFEYAVGSPMMHFSLTDNYV 220
 QY 356 VVPDQOVVF-----KLPEMT-----RGGS --PV 376
 Db 221 VIYDLPVTFDPNQVWVPPASVPRWLQRPARLVIQSVLGRVRFDPIAALGNRMQGHSDBLPY 280
 QY 377 YYDKNKVAFRGITLKDYAEDSSNIKWIDAPDCFCFHLYNAWEWPETD-EVYVIG---SCM 431
 Db 281 ANNPSPYARGVMPREG-GNEDYRWFDEIPECVYHPINAYSECRNGAEAVLVDDVVRYSRM 339
 QY 432 TPPDSIFNEDENLKSVLSEIRUNLKTCGESTRPRPISNEDQQVNLEAGMVNRMLGRKT 491
 Db 340 FDDRRGGDS -RPSDRWLNLTGSAVTAE-- -CRDDRA -QEFPRIETLVGGSPHR 392
 QY 492 FAY-----LALAEWPWKVSGFAKVYDLTGEVKHLGDNRVGGPLEPLPGE 538
 Db 393 FAYTYGEGFLYVGAGAALSTP-----LYKQDCVGSSTSAYSLDPPLIGMVFVPPNS 446

RESULT 14
 A:0582
 Hypothetical protein Rv0913c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Accession: A70582
 C:Accession: A70582
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Gentles, S.; Hamlin, N.; Holroyd, S.; Ratandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squires, S.; Nature 393, 537-544, 1998

Query Match	14.48;	Score 454.5;	DB 2;	Length 502;
Best Local Similarity	27.68;	Pred. No. 1.8e-26;		
Matches	139;	Conservative	73;	Mismatches 190; Indels 101; Gaps 18;

QY 14.1 WVGLLPDSIKGVYVRNGANPLHEPVTVGHHFFDGMVHAKVKEFHGSASYACRFQTQNRFY 200
 Db 41 VTEGPADDGILRLRNTPNPLHAFATVHPEFDGMIHVGFDDGKAFYRNFEIRTGDFL 100

QY 201 QEQRLGRPVPEPKAIGEHLHTGIGARLMLEYAR ---AAAGIVDPAHGV -ANAGLV ---252
 Db 101 AENPAGGPWLP-----GLAEPVQIAKREHGWARGLMDASSTDVTVRGIALTS 150

QY 253 -YFNGRLAMSEDDLPKQVQITPNGLKTVGRDFDGALESIMAHPKVDPSSGEALALS 311
 Db 151 FYQGDYRID ---PYSANL ---GKESWHGRPFEDPGVS ---AHKVDNNTGELLEFN 200

QY 312 YDVVKSYKLKYFRFSPDGTKSPDVEIQLDQPTMMHDFAITENFVPPDQVQVVFKLPEMIR 371
 Db 201 YS-KQEPMRYGVUDQNNELVHVDVLPDGPRLPHDMAFTENVILND-----247

QY 372 GSPPVVWDKKNVKAEGULDKY -----AEDSSNIKWDADPDCCEHUNNANEEPE 420
 Db 24.8 --FPLFWPRLLRBDVHLPRTFPEIPSPREAVARRGNDIRWEADPFEVLHTNAYE-Q 303

QY 421 TDEVVVTSGCMTPPSDNESD-----ENLKSVLSEFRNLNLKTE 460
 Db 304 GDEIVLVDG -----FYEGDQPLDGTGKWEKLFRLDRLQSRHLRORLNMTG- 353

QY 461 STRRPISSNEDQVN --LEAGMVNRNMLGRKTFKAFKLAEP-WPKVSGFAKVDLTGE 516
 Db 354 -----AVHEEQLSESSITEFGTSITNADYAASSYRTAATGRESWLFDFGLYVXKHDLLTG 406

QY 517 VKHHLGYNRYGEPLPGEGE -EDEGYLICFHDKEKTKWKSLSQIVNAVSLVEATVK 575
 Db 407 HECYSGFGVGVSETAMAPRVGSSAEDDGYLVLTDMNDASCYCLVEDAARPGDGPICK 466

QY 576 --LPSRVPYGFHCTFIGADDLAK 596
 Db 467 LALPERTSGGTISAWPAAELRR 489

RESULT 15
 S76169 hypothetical protein - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.

A:Variety: PCC 6803
 A:Accession: S76169
 C:Accession: S76169
 R:Kaneo, T.; Saito, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Yamada, M.; Ya. DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 A:Reference number: S74322; MUID:97061201; PMID:8905231

Result No.	Score	Query	Match	Length	DB ID	Description
1	120	3.8	616	1	NXFL_COTUJA	P58797 coturnix co
2	113	3.6	870	1	YCS0_YEAST	P25623 saccharomyc
3	113	3.6	466	1	DNAA_PROMI	P22837 proteus mir
4	107.5	3.4	658	1	PAK1_SCIBO	P50527 schizosacch
5	106	3.4	3562	1	PGCV_CHICK	Q90953 gallus gall
6	105	3.3	783	1	CAD5_MOUSE	P52284 mus musculu
7	105	3.3	2504	1	FAS_HUMAN	P49327 homo sapien
8	104.5	3.3	1079	1	AP0E_FREDDI	P16566 frenyella d
9	103	3.3	775	1	SYFB_AQUAE	P07620 aquifex aeo
10	102	3.2	2493	1	CYAA_USTMA	P49606 ustilago ma
11	101.5	3.2	520	1	GIGS_BRANA	Q9m462 brassica na
12	101.5	3.2	565	1	SCR1_SCIBO	Q14335 schizosacch
13	101	3.2	512	1	GUST_VICFA	P52417 vicia faba
14	100	3.2	3770	1	ACVS_EMENI	P27742 emericella
15	99.5	3.2	383	1	PHYC_BACSU	P31097 bacillus su
16	99.5	3.2	1256	1	ATL_STRAU	P52081 staphylococ
17	98	3.1	741	1	MASZ_MYCTU	Q50596 mycobacteri
18	98	3.1	1788	1	Y722_CAEFL	Q92921 caenorhabdi
19	97.5	3.1	518	1	CPT4_ARATH	Q96242 arabidopsis
20	97.5	3.1	691	1	SILP1_YEAST	P07955 saccharomyc
21	97	3.1	385	1	OXDC_BACSU	Q36714 bacillus su
22	97	3.1	636	1	KDPG_SCIBO	Q14019 schizosacch
23	96.5	3.1	426	1	UGR2_SCIBO	P78761 schizosacch
24	96.5	3.1	503	1	DLTA_BACSU	P39581 bacillus su
25	96.5	3.1	862	1	SMP2_YEAST	P32567 saccharomyc
26	96.5	3.1	3587	1	TYCB_BACBR	Q30408 b tyrocidin
27	96	3.0	462	1	DNAA_YERINIA	Q829u yersinia pe
28	95.5	3.0	865	1	NRFA_PINUR	Q92269 penicillium
29	95.5	3.0	918	1	DNUL_RATM	Q9jhy8 rattus norv
30	95.5	3.0	1144	1	DP3A_NEMIB	Q9Jxz2 neisseria m
31	94.5	3.0	399	1	EX7L_CLOAB	Q97hd0 clostridium
32	94.5	3.0	590	1	HMAA_DROME	P29555 drosophila
33	94.5	3.0	638	1	GHR_RAT	P16310 rattus norv

ALIGNMENTS						
<hr/>						
Scoring table:	BLOSUM62	Gapop 10.0	Gapext 0.5			
Searched:	112892 seqs, 41476328 residues					
Total number of hits satisfying chosen parameters:	112892					
Minimum DB seq length: 0						
Maximum DB seq length: 2000000000						
Post-processing: Minimum Match 0%						
Maximum Match 100%						
Database :	SwissProt;40.0*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
<hr/>						
SUMMARIES						
<hr/>						
RESULT 1						
ID NXFL_COTUJA						
STANDARD;						
PRP; 616 AA.						
RESULT FROM N.A., AND MUTAGENESIS OF GLU-246.						
RP MEDLINE:21257892; Published:11358B04;						
RA Cogburn G.A., Wiegand H.L., Kang Y., Ho D.N., Georgiadis M.M.,						
RA Cullen B.R.; RT using viral species specificity to define a critical protein/RNA interaction surface;"						
RT Genes Dev. 15:1194-1205(2001).						
RN [1]						
CC FUNCTION: Involved in the nuclear export of mRNA from the nucleus						
CC to the cytoplasm.						
CC [-] SUBCELLULAR LOCATION: Nuclear.						
CC [-] DOMAIN: The leucine-rich repeats and the NTF2-domain are essential for the export of mRNA from the nucleus (BY similarity).						
CC [-] MISCLENEOUS: Changing a single residue, Glu-246 to Arg, the residue found in human NXFL, fully rescue both constitutive transport element (CTE) function and CTE binding.						
CC [-] SIMILARITY: BELONGS TO THE NXFL FAMILY.						
CC [-] SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).						
CC [-] SIMILARITY: CONTAINS 1 NTF2 DOMAIN.						
CC [-] SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).						
CC [-] SIMILARITY: CONTAINS 1 QBA DOMAIN.						
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).						
DR EMBL: AF343749; AAK56910.1; -						
DR PROTEIN: PS50177; NTF2_DOMAIN; 1.						
KW Transport; mRNA transport; Nuclear protein; Repeat;						
KW Leucine-rich repeat; Multigene family.						
FT DOMAIN 116 195 RNA-BINDING (RRM) (BY SIMILARITY).						
FT REPEAT 263 288 LRR 1.						
FT REPEAT 289 312 LRR 2.						
FT REPEAT 313 340 LRR 3.						
FT REPEAT 341 368 LRR 4.						
FT DOMAIN 383 533 NTF2.						
FT DOMAIN 563 605 UBA-LIKE (BY SIMILARITY).						

FT	MUTAGEN	246	246	O>RESUES CTE FUNCTION AND BINDING.
SQ	SEQUENCE	616 AA;	69329 MW;	D14903DDCF1BF1BD CRC64;
Query Match	Best Local Similarity	3.8%	Score 120;	DB 1; Length 616;
Matches	Local Similarity	22.0%	Pred. No 0.29;	Gap
86	PSRAPIVTRDLPPIPDGGSSRDGRRNWKITIPIYGKKYRTWLLSSIQNCSVPPT	1	1	1
86	PSRAPIVTRDLPPIPDGGSSRDGRRNWKITIPIYGKKYRTWLLSSIQNCSVPPT	1	1	1
329	GTKSPVEIQLDQPT	1	1	1
329	GTKSPVEIQLDQPT	1	1	1
282	GRFDGQGLESTM	1	1	1
28	GRGPFRGKMSMNRNSRNRRGGTGPSSSRQRQEDGDV-AMS DAHDAPRGRRLPYGPR	8	1	1
QY	QVVFKLPEMIRGGSPVYDKNNKVAFRGFLDKYAEDESSNIKWIDAPDCFCFHLLNAWEEPE	4	1	1
QY	QVVFKLPEMIRGGSPVYDKNNKVAFRGFLDKYAEDESSNIKWIDAPDCFCFHLLNAWEEPE	4	1	1
Db	PVEFH	1	1	1
Db	PVEFH	1	1	1
QY	TDEVVYIGSCMTPPDSIFN	1	1	1
QY	TDEVVYIGSCMTPPDSIFN	1	1	1
Db	YDHNR-AQFYVED	1	1	1
Db	YDHNR-AQFYVED	1	1	1
QY	ATTASAIKQS	1	1	1
QY	ATTASAIKQS	1	1	1
146	EIRLNLKIGESTRPIISN	4	1	1
146	EIRLNLKIGESTRPIISN	4	1	1
421	EDQVNUNLEAGMVNRNLGRKTKEFAYLALAEPWPVKSGEAKVD	5	1	1
421	EDQVNUNLEAGMVNRNLGRKTKEFAYLALAEPWPVKSGEAKVD	5	1	1
181	NYKVVIIINSSAPPSSIQNELKPEE1EQLVCMISKRYDGAQRDLK GLRVDPDLVSOS	2	1	1
470	EDQVNUNLEAGMVNRNLGRKTKEFAYLALAEPWPVKSGEAKVD	5	1	1
240	IDVVLNQRSCMMVVLRIIEENIPELOSLNLSNMKNLYKLDDLAEL	1	1	1
240	IDVVLNQRSCMMVVLRIIEENIPELOSLNLSNMKNLYKLDDLAEL	1	1	1
Qy	LTTGGVKKHLGYDNRGY--GFPFLPFG--GGEDEGSYILCFVHDEKTWKSSL--QIV	5	1	1
Qy	LTTGGVKKHLGYDNRGY--GFPFLPFG--GGEDEGSYILCFVHDEKTWKSSL--QIV	5	1	1
295	LSRNELKSERBLDKYGLKLBELWLGNPLCDQFDDOSSYISVRREPPLRFLDGHFLP	35	1	1
564	NAVSLEVEATVKLPSRVYPYGFHGTFGADDL	5	1	1
355	PP1AFDVEAPVTLPP	1	1	1
Db	-CKGSYFGSDDL	3	1	1
Db	-CKGSYFGSDDL	3	1	1
RESULT 2				
YCS0_YEAST	ID YCS0_YEAST STANDARD;	PRT;	870 AA.	
AC	P25673; P25622;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	15-DBC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical protein in RIMI-RPS14A intergenic region.			
GN	YCR030_C_R YCR30C/YCR29C			
OS	Saccharomyces cerevisiae (Baker's Yeast)			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
CX	NCBI_TAXID=4942;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RQ	Cederbaum R., Hohmann S., Schaaff-Gerstenschaeger I., Huse K., Zimmermann F.K.;			
RQ	Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.			
RQ	[2]			
RQ	REVISIONS.			
RA	Gromadka R.; Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.			
RQ	[1]			
RQ	SEQUENCE FROM N.A.			
RA	Cederbaum R., Hohmann S., Schaaff-Gerstenschaeger I., Huse K., Zimmermann F.K.;			
RA	Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.			
RQ	[2]			
CC	SIMILARITY: SOME, TO S.POMBE SPBC4C3_06.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcements.html or send an email to license@isb-sib.ch).			
CC	EMBL; X59720; CAA47322_1; -			
DR	SCD; S000626; YCR030C.			
KW	Hypothetical protein.			
SQ	SEQUENCE 870 AA;	95125 MW;	3F2CE2F1562E5277 CRC64;	

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

RT Cdc42p and is involved in the control of cell polarity and mating.
 RL EMBO J. 14: 5908-5919(1995).
 RN [2]
 RP SEQUENCE FROM N.A.

RT Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
 RL [3]
 RN [3]
 RP SEQUENCE FROM N.A.
 RP STRAIN=972;
 RX MEDLINE=21BA48401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lycett R., Stewart A.,
 RA Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Harris N., Hodgetts D., Hidalgo J., Hunt S., Jacobs K.,
 RA Holloway S., Hornsby T., Howarth S., Huckle B.J., Jones M., James K.,
 RA Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mangall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbaniwitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkert G., Aert J., Robben J., Grymonprez B.,
 RA Weltjens I., Vansteels E., Schaefer M., Mueller-Puer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Eber P., Zimmerman N., Weidler H., Wambutt R., Purcellie B.,
 RA Gaffreau A., Cadieu E., Dreano S., Gloux S., Lealauri V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe.",
 RL Nature 415:871-880 (2002).
 RN [4]

RT SEQUENCE OF 119-658 FROM N.A.
 RX MEDLINE=95340235; PubMed=597098;

QY 105 VSHEKHPPLPKT-ADPSVQIAGNAPVNPFVRRNLPPVGLPDS-TKGYYVRNGANPLH 21;
 Db 83 VSARTESVPKTVTTHAV---NSTPTNSQPYR--PSWDNQPQSQLEPNYRSNNPKH 135
 QY 163 EPVTGHFFPDGMVHAKFVEHGSASYACRFETOTNRVFOERQLG--RPVPEKAIGELGH 220
 Db 136 K-----FD-----NFVEGKSQNLAR -AAAQQVADNPGGAYNPLF-----LYGG 172
 QY 221 TGIAIRMLFYARAAGIVDPAHGTG-----ANAGLYVFNGRLLANSEDDLQPYQVLTPN 276
 Db 173 TGLGKTHLLH-----AVGNSIMERKANAKVYMH-----SERPVQDMYKALQNN 216
 QY 277 DLKTVGRFDQGLESTMIAHPKVDPESEGLFALSYSTDVVKSPYKLFRSPDGTKSPDVE 336
 Db 217 AIE-----DFK-----RYR-SYDALLDDIQ 237
 QY 337 IQLDQPTMMHDFAITENFVVVDPQQVVF---KLPEMTRGGSPVWYDKNKYVAFRGTLDKYA 393
 Db 238 FFANKERSQEEFFHFFNALLLEGNOQTILISDRYKEITNG----VERLRK-SRFG----- 286
 QY 394 EDSSNWKWIDAPDCFCFHLMNAWEEPETDVYVIGSCMTTPDSIENESDENLKSVLSEI- 452
 Db 287 -----WGLTAIEPEPELETFRAI-----LMKKADENQIOLPDEVA 321
 QY 453 -----RLNIKTAGESTRRPILSN 469
 Db 322 FFIAKRLRSNVRLEGALNRYIAN 345
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN_KINASES.
 CC -!- SIMILARITY: CONTAINS 1 CRIB DOMAIN.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

CC DR U22371; AAC49105.1;
 CC DR EMBL; AL034433; CAA223347.1; -;
 CC DR HSSP; Q00534; IBI8;
 CC DR InterPro; IPR000719; Euk_kinase;
 CC DR InterPro; IPR00095; PAKbox/RhoBndng;
 CC DR InterPro; IPR002290;
 CC DR InterPro; IPR001245; Tyr_kinase;
 CC DR Pfam; PF00069; Pkinase; 1;
 CC DR PRINTER; PR0109; TYR_KINASE;
 CC DR ProdDom; PDD00001; Euk_kinase; 1.

RESULT 4

ID	PAK1_SCHPO	STANDARD:	PRT:	658 AA.
AC	P50527;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Serine/threonine-protein kinase paki1/shk1 (EC 2.7.1.-).			
GN	PAK1 OR SHK1 OR ORB2 OR SPBC1G04.1.c.			
OS	Schizosaccharomyces pombe (Fission yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetidae;			
OC	Schizosaccharomycetes; Schizosaccharomyctaceae;			
OX	Schizosaccharomycetes.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96112805; PubMed=8846783;			
RA	Ottlie S., Miller P.J., Johnson D.I., Creasy C.L., Sells M.A.,			
RA	Bagrodia S., Forsburg S.L., Chernoff J.J.			
RT	"Fission yeast PAK1 encodes a protein kinase that interacts with			

DR SMART; SM00285; PBD; 1.	RN SEQUENCE FROM N.A.
DR PROSITE; PS50108; CRIB; 1.	RP STRAIN=White Leghorn; TISSUE=Limb bud;
DR PROSITE; PS50107; PROTEIN_KINASE_ATP; 1.	RC MEDLINE=3300846; PubMed=8314802;
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.	RX Shionomura T., Nishida Y., Ito K., Kimata K.;
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.	RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed during chondrogenesis in chick limb buds. Alternative spliced multi-forms of PG-M and their relationships to versican,";
KW Transfase; Serine/threonine-protein kinase; ATP-binding; phosphorylation.	RT J. Biol. Chem. 268:14461-14469(1993).
FT DOMAIN 147 160 CRIB.	RL -!- FUNCTION: May play a role in intercellular signaling and in connecting cells with the extracellular matrix. May take part in the regulation of cell motility, growth and differentiation. Binds hyaluronic acid.
FT DOMAIN 85 88 POLY-SER.	CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
FT DOMAIN 234 237 POLY-SER.	CC -!- ALTERNATIVE PRODUCTS: At least 2 isoforms; V0 (shown here) and V1; are produced by alternative splicing.
FT DOMAIN 246 249 POLY-SER.	CC -!- TISSUE SPECIFICITY: Prechondrogenic condensation area of developing limb buds.
FT DOMAIN 263 267 POLY-SER.	CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development (by similarity).
FT DOMAIN 386 637 PROTEIN_KINASE.	CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
FT NP_BIND 392 400 ATP (BY SIMILARITY).	CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
FT BINDING 415 415 ATP (BY SIMILARITY).	CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
FT ACT_SITE 505 505 BY SIMILARITY.	CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
FT CONFLICT 492 495 GLOH -> LYSD (IN REF. 1).	CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
FT CONFLICT 537 537 R -> P (IN REF. 1).	CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announce/ or send an email to license@ebi.ac.uk).
SEQUENCE 658 AA: 72358 MW: 697255C925021E5 CRC64;	CC -!- SIMILARITY: DB 1: IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
Query Match 3.4% Score 107.5; DB 1; Length 658; Best Local Similarity 20.6%; Pred. No. 2.7%; Matches 88; Conservative 61; Mismatches 157; Indels 121; Gaps 20;	CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
Qy 18 NHTQPLSSQQSSL-----SYCSSLEMASRVRKLNVSSAHLTPPAHLFPKQSS 67	CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
Db 237 SHLQPTATSSRLYPSRSPRTPASSSSQVTTITNSRQPSPLVSSKSTD 296	CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
Qy 68 N----SPATIVVKPKRENSNTKQM---NLVQRAAAALDAAEGLVYSHKEKLHPKPTA 117	CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
Db 297 NTRSHSPVLLTPQTLSSTSEKPHIRPNNTSIVQRRAETSTK-----PKAV 341	CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
Qy 118 DESVQLAGNFIAVNEQFVRNLPVYVKGKLPDSIKGVYVRNGANPLHEPVTCIHFEDGDMV 177	CC -!- SIMILARITY: DB 1: IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
Db 342 ATPQKYEAPSARLQLQRAPRO----QSNDNAVLAXLQSTCNPKNPTLLRN----- 389	CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
Qy 178 HAVKEFHGSAS-YACRFQTNRFVQEROGLCPVFK--AIGEL----HGHTGTARLM 227	CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
Db 390 -TVKIGQASSEDVYSRQVGTNLSAIKKNNIQQPKEFVNELLVMKSHHHKIVNF 447	DR InterPro: IPR000152; Asx_hydroxy.
Qy 228 -LFYARAAGIIVDPAHGTG--VANAGLYFNGRNLAMEDDPYQVOITPNG---D 277	DR InterPro: IPR000742; EGF_2.
Db 448 DFFYKSELNMMEYMRGGSLTEVYVNTNL--SEQQIAAICKETEGLQLHENGIVHRD 505	DR InterPro: IPR001881; EGF_Ca.
Qy 278 LKT-----VGREFDDQLES-----TMIAHP-----KVDPES 304	DR InterPro: IPR003599; Ig.
Db 506 IKSDNLLSLQDILTDIFSCFAQIDSNTMKTRKTRNNGTPWMAPVVTRKEYGFYDVWS 565	DR InterPro: IPR003006; Ig_MHC.
Qy 305 GELPALSYDVSQSKPYL-----YFRFSPDGTKSPDVTEQDQPTMM----HDFAITENF 354	DR InterPro: IPR001304; Lectin_C.
Db 566 LGIMIALEMVEGPYPNLNPRLRAYLIATGP-----KISRPLLSSYHDF-LSKSL 618	DR InterPro: IPR000538; Link.
Qy 355 VVVPDQO 361	DR Pfam: PF000049; EGFR.
Db 619 TNPKQR 625	DR Pfam: PF000436; Sushi_SCR_CCP.
RESULT 5	DR Pfam: PF000068; EGFR.
PGCV_CHICK	DR Pfam: PF00047; Ig.
ID PGCV CHICK STANDARD; AC Q9053; Q90945;	DR Pfam: PF00059; lectin_c.
DT 01-NOV-1997 (Ref. 35, Created)	DR Pfam: PF00084; sushi.
DT 15-JUN-2002 (Ref. 41, Last sequence update)	DR SMART: SM00032; CCP.
DE Versican core protein precursor (large fibroblast update)	DR SMART: SM00034; CLECT.
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).	DR SMART: SM00179; EGF_Like.
GN CPSG2.	DR SMART: SM00041; EGF_Like.
OS Gallus gallus (Chicken)	DR SMART: SM00409; Ig.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Sushi; Gallus. NCBI_TAXID=9031;	DR SMART: SM00445; Link.
FT SIGNAL 1	DR PROSITE: PS00010; ASX_HYDROXYL; 1.
FT CHAIN 27	DR PROSITE: PS00022; EGF_1; 2.
FT DOMAIN 37	DR PROSITE: PS01186; EGF_2; 1.
FT DOMAIN 136	DR PROSITE: PS01187; EGF_Ca; 1.
FT DOMAIN 243	DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
FT DOMAIN 166	DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
FT DOMAIN 264	KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi; Signal; Repeat; EGFR-like domain; Calcium; Immunoglobulin domain; Hyaluronic acid; Alternative splicing.
FT DOMAIN 26	KW POTENTIAL.
FT DOMAIN 3562	KW VERSICAN CORE PROTEIN.
FT DOMAIN 136	KW IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 243	KW LINK 1.
FT DOMAIN 345	KW LINK 2.

DOMAIN	3254	3290	EGF-LIKE 1.	Score: 106;	PDB ID:	Db	1496	--QQVTSAPKSTKAEEARSDD--QIES--VAHSK--NVTFSQINENTNFELSET--
DOMAIN	3292	3328	EGF-LIKE 2.	Score: 106;	DB ID:	Qy	322	YFRFSSDGTKSPDV----ETQLODQPTMMHDEATENVVVFDQVVFKLPEMIGGSP
DOMAIN	3341	3455	C-TYPE LECTIN.	Score: 106;	DB ID:	Db	1541	--NIKWIDAPDCFCFHWNAAWE----PETDEV
DOMAIN	3460	3518	SUSHI.	Score: 106;	DB ID:	Qy	376	VYDKNKVARGFLDKYEDSS-----
DISULFITID	44	129	BY SIMILARITY.	Score: 106;	DB ID:	Db	1573	WYSGESEVT--TTDKYLEITSVYEQSPKKNKETV-----MWHTGEESSTKTKNL
DISULFITID	171	216	BY SIMILARITY.	Score: 106;	DB ID:	Qy	425	WVIGSCMTPPDSIENSDENLKSVLSEIRLNLTGESTRR
DISULFITID	195	344	BY SIMILARITY.	Score: 106;	DB ID:	Db	1621	LLITNESSGDGS--TSDLS-RSFETI-LTMSSHEDSEK
DISULFITID	269	314	BY SIMILARITY.	Score: 106;	DB ID:	Qy	44	QVTSAPKSTKAEEARSDD--QIES--VAHSK--NVTFSQINENTNFELSET--
DISULFITID	293	314	BY SIMILARITY.	Score: 106;	DB ID:	Db	1573	WYSGESEVT--TTDKYLEITSVYEQSPKKNKETV-----MWHTGEESSTKTKNL
DISULFITID	3258	3269	BY SIMILARITY.	Score: 106;	DB ID:	Qy	425	WVIGSCMTPPDSIENSDENLKSVLSEIRLNLTGESTRR
DISULFITID	3263	3278	BY SIMILARITY.	Score: 106;	DB ID:	Db	1621	LLITNESSGDGS--TSDLS-RSFETI-LTMSSHEDSEK
DISULFITID	3280	3289	BY SIMILARITY.	Score: 106;	DB ID:	Qy	44	QVTSAPKSTKAEEARSDD--QIES--VAHSK--NVTFSQINENTNFELSET--
DISULFITID	3292	3307	BY SIMILARITY.	Score: 106;	DB ID:	Db	1573	WYSGESEVT--TTDKYLEITSVYEQSPKKNKETV-----MWHTGEESSTKTKNL
DISULFITID	3301	3316	BY SIMILARITY.	Score: 106;	DB ID:	Qy	425	WVIGSCMTPPDSIENSDENLKSVLSEIRLNLTGESTRR
DISULFITID	3318	3327	BY SIMILARITY.	Score: 106;	DB ID:	Db	1621	LLITNESSGDGS--TSDLS-RSFETI-LTMSSHEDSEK
DISULFITID	3334	3345	BY SIMILARITY.	Score: 106;	DB ID:	Qy	44	QVTSAPKSTKAEEARSDD--QIES--VAHSK--NVTFSQINENTNFELSET--
DISULFITID	3362	3454	BY SIMILARITY.	Score: 106;	DB ID:	Db	1573	WYSGESEVT--TTDKYLEITSVYEQSPKKNKETV-----MWHTGEESSTKTKNL
DISULFITID	3430	3446	BY SIMILARITY.	Score: 106;	DB ID:	Qy	425	WVIGSCMTPPDSIENSDENLKSVLSEIRLNLTGESTRR
DISULFITID	3461	3504	BY SIMILARITY.	Score: 106;	DB ID:	Db	1621	LLITNESSGDGS--TSDLS-RSFETI-LTMSSHEDSEK
DISULFITID	3490	3517	BY SIMILARITY.	Score: 106;	DB ID:	Qy	44	QVTSAPKSTKAEEARSDD--QIES--VAHSK--NVTFSQINENTNFELSET--
CARBOYD	163	163	(POTENTIAL).	Score: 106;	DB ID:	AC	P552B4; O35542;	
CARBOYD	235	235	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	DT	O1-OCT-'1996 (Rel. 34. Created)	
CARBOYD	329	329	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	DT	O1-OCT-'1996 (Rel. 40. Last sequence update)	
CARBOYD	529	529	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	DT	16-OCT-2001 (Rel. 40. Last annotation update)	
CARBOYD	709	709	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	DE	Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5).	
CARBOYD	948	948	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	GN	CDH5.	
CARBOYD	1409	1409	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	OS	Mus musculus (Mouse)	
CARBOYD	1479	1479	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CARBOYD	1523	1523	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	OC	Mammalia; Butheria; Rodentia; Sciurognathia; Muridae; Mus.	
CARBOYD	1530	1530	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	OX	NCBI_TaxID=10090;	
CARBOYD	1625	1625	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RN	[1]	
CARBOYD	1751	1751	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RP	SEQUENCE FROM N.A., TISSUE: Brain capillary;	
CARBOYD	1988	1988	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RX	MEDLINE=9641083; PubMed=8555485;	
CARBOYD	2019	2019	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RA	Breiter G., Caveda L., Berthier R., Schnuerch H.,	
CARBOYD	2088	2088	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RA	Gotsch U., Vestreboer D., Risau W., Dejana E.;	
CARBOYD	2089	2089	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RT	"Molecular cloning and expression of murine vascular endothelial-cadherin in early stage development of cardiovascular system."	
CARBOYD	2507	2507	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RT	Blood 87:630-641(1996).	
CARBOYD	2642	2642	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RL	RT	
CARBOYD	2679	2679	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RN	SEQUENCE FROM N.A., AND FUNCTION;	
CARBOYD	2748	2748	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RP	TISSUE=BALB/C; TISSUE: Breast carcinoma;	
CARBOYD	2762	2762	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RX	MEDLINE=97364256; PubMed=9220534;	
CARBOYD	3069	3069	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RX	Takeichi M., Imamura S.;	
CARBOYD	3194	3194	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RA	Matsuishi N., Toda K.-I., Horiguchi Y., Tanaka T., Nakagawa S.;	
CARBOYD	3232	3232	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RA	"In vivo evidence of the critical role of cadherin-5 in murine vascular integrity."	
CARBOYD	3545	3545	N-LINKED (GLCNAC. . .)	Score: 106;	DB ID:	RT	Proc. Assoc. Am. Physicians 109:362-371(1997).	
VARSPIC	485	1411	MISSING (IN ISOFORM VI).	Score: 106;	DB ID:	RL	-1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.	
SEQUENCE	3562	AA;	38B078 MW:	Score: 106;	DB ID:	CC	THEY PREFERentially INTERACT WITH THEMSELVES IN A HOMOPHILIC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE SORTING OF HETEROGENEOUS CELL TYPES THIS CADHERIN MAY PLAY A	
Q	Query Match	3.4%	Score: 106;	DB ID:	CC	IMPORTANT ROLE IN ENDOTHELIAL CELL BIOLOGY THROUGH CONTROL OF THE COHESION AND ORGANIZATION OF THE INTERCELLULAR JUNCTIONS.		
Best Local Similarity	22.7%	Pred. No. 41;	No. 41;	Length 356;	DB ID:	CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.	
Matches 118: Conservative	61:	Mismatches 191:	Indels 150:	Gaps 29;	DB ID:	CC	-1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.	
Matches	118:				CC	-		
1	MASFWATA-AVSGRWRGNNHTOPPLSSQDSSDLYCSSLPLMASRVTRKLNVSSALHT--	56	DB ID:	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this version is not removed. Usage by commercial entities requires a license agreement (See http://www.isb-sib.ch/announcements or send an email to license@isb-sib.ch).			
1231	MISSPFTAGSLITLGASPTPEGSSISLEEVKTVPSRATVIVTIVSDFLSSI	1290	DB ID:	CC	-			
57	-PPALHFPKOSSNSPAVIVKEKAKESNTKQNL----FQRAAAALDAAEFL	104	DB ID:	DR	EMBL; D63940; CAA58789; 1-			
1291	SAVDK1QPTSKASKPVSFKSSPSKDIVIDESIPTVSSDADDLTKM	1348	DB ID:	DR	MGD; MGI:105057; Cdhs5;			
105	VSHEKHLPLK-----TADPSIQAGN FAPVNEQPVRNLPVFGKLPS-	148	DB ID:	DR	InterPro; IPR002126; Cadherin.			
1349	VEPE----IDEKEYFTSSTATAVRPTAPPTMEATEALQPEVSPSH-----	111	DB ID:	DR	InterPro; IPR00233; Cadherin_C-term.			
1411	MISSING (IN ISOFORM VI).	111	DB ID:	DR	PFam; PF00128; Cadherin_5;			
1457	YIILDFFP-----NFMDFEEE--OCENTADYVTPALQFLNGK-----	1495	DB ID:	DR	PFam; PF010149; Cadherin_C-term; 1.			
262	SEDDPLQVQIPNGLDTKVGREDDFQLESIMIAHPVDPESGSELFASYDVKSPYLR	321	DB ID:	DR	SMART; SMART112; CA_5;			
Q	Query Match	3.4%	Score: 106;	DB ID:	DR	SMART; SMART112; CA_5;		
Best Local Similarity	22.7%	Pred. No. 41;	No. 41;	Length 356;	DB ID:	DR	YIILDFFP-----NFMDFEEE--OCENTADYVTPALQFLNGK-----	
Matches 118: Conservative	61:	Mismatches 191:	Indels 150:	Gaps 29;	DB ID:	DR	PFam; PF00128; Cadherin_5;	
Matches	118:				DB ID:	DR	SEDDPLQVQIPNGLDTKVGREDDFQLESIMIAHPVDPESGSELFASYDVKSPYLR	

DR	PROSITE; PS00232; CADHERIN_1; 3.	Db	647 EGGEMDFTSYDVSLNNSVRGGSTKPLRSTMDARPATVQVKPRLAPGLHGS---SRE 703
DR	PROSITE; PS50268; CADHERIN_2; 5.	QY	594 LAKQV 598
KW	Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;	Db	704 MATMI 708
KW			
WT	SIGNAL .		
WT	1 24	POTENTIAL .	
WT	PROPEP 25 45	POTENTIAL .	
WT	CHAIN 46 783	VASCULAR ENDOTHELIAL-CADHERIN .	
WT	DOMAIN 46 592	EXTRACELLULAR (POTENTIAL) .	
WT	TRANSMEM 593 619	POTENTIAL .	
WT	DOMAIN 620 783	CYTOSMERIC (POTENTIAL) .	
WT	DOMAIN 46 148	CADHERIN 1 .	
WT	DOMAIN 149 255	CADHERIN 2 .	
WT	DOMAIN 256 370	CADHERIN 3 .	
WT	DOMAIN 371 475	CADHERIN 4 .	
WT	DOMAIN 476 592	CADHERIN 5 .	
WT	DOMAIN 737 752	SER-RICH .	
WT	CARBOHYD 59 59	N-LINKED (GLCNAC . . .) (POTENTIAL) .	
WT	CARBOHYD 154 154	N-LINKED (GLCNAC . . .) (POTENTIAL) .	
WT	CARBOHYD 440 440	N-LINKED (GLCNAC . . .) (POTENTIAL) .	
WT	CARBOHYD 522 522	N-LINKED (GLCNAC . . .) (POTENTIAL) .	
WT	CARBOHYD 534 534	N-LINKED (GLCNAC . . .) (POTENTIAL) .	
WT	67 KQQ -> GKI (CIN REF. 2).	KQQ -> GKI (CIN REF. 2).	
SQ	SEQUENCE 783 AA; 87847 MW; 240AFD663BCE71C CRC64;	RN [1].	
		RP SEQUENCE FROM N.A.	
		RC TISSUE=brain	
		RX MEDLINE=96004605; PubMed=7567999.	
		RA Jayakumar A.; Tai M.-H.; Huang W.-Y.; Al-Feel W.; Hsu M.,	
		RA Abu-Elheiga L.; Chirala S.S.; Wakil S.J.;	
		RT "Human fatty acid synthase: properties and molecular cloning.";	
		RL Proc. Natl. Acad. Sci. U.S.A. 92:8695-8699(1995).	
		RN [2].	
		RP SEQUENCE OF 753-758 AND 1285-1297.	
		RX MEDLINE=94293385; PubMed=8022791.	
		RA Kujada F.P.; Jenner J.; Wood F.D.; Hennigar R.A.; Jacobs L.B.,	
		RA Dick J.D.; Pasternack G.R.,	
		RT "Fatty acid synthesis: a potential selective target for	
		RT anticancer therapy";	
		CC -!- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF	
		CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.	
		CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN	
		ACYL CARRIER PROTEIN.	
		CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a	
		CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).	
		CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = COA +	
		CC acetyl-[acyl-carrier protein]; Malonyl-CoA + [acyl-carrier protein] = COA +	
		CC malonyl-[acyl-carrier protein].	
		CC -!- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-	
		CC carrier protein] = -oxoacyl-[acyl-carrier protein] + CoA +	
		CC [acyl-carrier protein].	
		CC -!- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl]-carrier protein] +	
		CC NADP(+) = 3'-oxoacyl-[acyl-carrier protein] + NADPH.	
		CC -!- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl]-carrier protein]	
		CC = 2-hexadecenoyl-acyl-carrier protein + H(2)O.	
		CC -!- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-	
		CC -!- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + NADPH.	
		CC -!- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-	
		CC carrier protein] + oleate.	
		CC -!- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.	
		CC -!- TISSUE SPECIFICITY: PROMINENT EXPRESSION IN BRAIN, LUNG, AND	
		LIVER.	
		CC -!- MISCELLANEOUS: THE RELATIVELY LOW BETA-KETOACYL SYNTHASE ACTIVITY	
		CC MAY BE ATTRIBUTABLE TO THE LOW 4'-PHOSPHORANTHEINE CONTENT OF	
		THE PROTEIN.	
		CC -----	
		This SWISS-PROT entry is copyright. It is produced through a collaboration	
		CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	
		CC the European Bioinformatics Institute. There are no restrictions on its	
		use by non profit institutions as long as its content is in no way	
		modified and this statement is not removed. Usage by and for commercial	
		entities requires a license agreement (See http://www.isb-sib.ch/announce/	
		or send an email to license@isb-sib.ch).	

Query Match Similarity Score 103; DB 1; Length 775;
 Best Local Similarity 20.7%; Pred. No. 7.4; Mismatches 67; Indels 212; Gaps 31;
 Matches 114; Conservative 114;

QY 158 ANPLHPPVTGHFFDGDM-VHAKPEHGSAS-----YACRET 194
 Db 555 -----VFGDRDVYESQRLYQEIKENQLSVREFIRALAKSDVERKTYWTSLYVCKAI 607
 DR 555 -----VFGDRDVYESQRLYQEIKENQLSVREFIRALAKSDVERKTYWTSLYVCKAI 607

QY 195 QINRFVQERQLGRPVKEPKATGELHGTGIAIRMLFYARAAGTIVDPAHGTGVANGLVYF 254
 Db 608 E---YIHRRLGRPTGRO-EINKYFDIAAKQGFZA----VVV-----AIVNS--VEY 650
 DR 608 E---YIHRRLGRPTGRO-EINKYFDIAAKQGFZA----VVV-----AIVNS--VEY 650

QY 255 NGRLAMSEDDLPYQVQITPN---DLKTVG--RDFDGGLESTIMIAHPKVDPESGEFL 308
 Db 651 SE--AFGEDTPYERYLTPSGVALQRQ-VGSIREDVVG----VYQKQETPLF 696
 DR 651 SE--AFGEDTPYERYLTPSGVALQRQ-VGSIREDVVG----VYQKQETPLF 696

QY 309 ALSYDVSXPKYLYFRSPDKPSDKTPSKPDEVIQDQ 341
 Db 697 VFLGTVT-----DIRTEPDIQFRINQ 717
 DR 697 VFLGTVT-----DIRTEPDIQFRINQ 717

RESULT 9

SYTB_AQUAE STANDARD; PRT; 775 AA.
 AC 061620;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE tRNA 1-glycosyltransferase beta chain (EC 6.1.1.20) (Phenylalanine--tRNA 1-glycosyltransferase chain) (PfRS).
 GN PHET OR AQ_1730.
 OS Aquificae; Aquifacae (class); Aquificales; Aquificaceae;
 OC Aquificae; Aquifacae (class); Aquificales; Aquificaceae;
 CX NCB_TaxID=63363;
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=V15;
 RX MEDLINE=88196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Shadid M.A., Keller M., AuJay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus";
 RL Nature 394:353-358 (1998).
 CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
 CC diphosphate + L-phenylalanyl-tRNA(Phe).
 CC -!- SUBSTRATE: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
 CC CHAIN FAMILY. SUBFAMILY 1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AE000753; AAC07582.1; -.
 DR HSSP; P27002; IEPYS.
 DR InterPro; IPR005146; B3_4.
 DR InterPro; IPR005147; B5.
 DR InterPro; IPR005121; Fdx-AntiCB.
 DR InterPro; IPR004522; Phet_bact.
 DR Pfam; PF01588; tRNA_bind_1.
 DR Pfam; PF03147; EDX-ACB; 1.
 DR Pfam; PF03483; B3_4; 1.
 DR Pfam; PF03484; B5_1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase: ATP-binding.
 KW Complete proteome. 775 AA;
 KW Sequence 775 AA; 88484 MW; 61BB15FDE5B274C8 CRC64;

Query Match Similarity Score 103; DB 1; Length 775;
 Best Local Similarity 20.7%; Pred. No. 7.4; Mismatches 67; Indels 212; Gaps 31;
 Matches 114; Conservative 114;

QY 123 IAGNFAPVNQSPVRRNLPPVVKLPPDSIKGYVVRNGANPLHPPVTGHFFD---- 172
 Db 46 VEGKVVVEKEHPTKKLAVV-----KVOQEHIIDVVTVDKSRE 86
 DR 46 VEGKVVVEKEHPTKKLAVV-----KVOQEHIIDVVTVDKSRE 86

QY 173 GDGMVHAYKFEHGSASYACRETQTNRFQEROLGRPVPKAIGEZHHTGARLMFYAR 232
 Db 87 GDVVI--VALPAKVGNCM-----VTERE-----DGVVSKGILLSA 121
 DR 87 GDVVI--VALPAKVGNCM-----VTERE-----DGVVSKGILLSA 121

QY 233 AAAGIVDAHGTGVANAGLVYFNGRLLAMSEDDLP-----YQQTIPN-GD 277
 Db 122 QELGLEPKSEG-----VLIKIHEPDFKPQGDANEITLGFFERKIEIDITNRGD 167
 DR 122 QELGLEPKSEG-----VLIKIHEPDFKPQGDANEITLGFFERKIEIDITNRGD 167

QY 278 LKTVGRPFDFGQLESTMIAHPKVDP----BSGELFALSVDVSKPVLKYFRFSPDGTK 331
 Db 168 MLSVRGYARD--LSAIFRLPKKKPEEPTYEETGEFFITEDEDCKRY-----RGVV 216
 DR 168 MLSVRGYARD--LSAIFRLPKKKPEEPTYEETGEFFITEDEDCKRY-----RGVV 216

QY 332 SPVYEIQDQP-----TMMHDFATENFVVPDQSVVFVKLPEMRGGSPV-VY 378
 Db 217 IEQVEIK-ESPLYIKKRIWQCGIKSINNNVDTINYV-----MLRDGQPHAF 262
 DR 217 IEQVEIK-ESPLYIKKRIWQCGIKSINNNVDTINYV-----MLRDGQPHAF 262

QY 379 DKQKVARYGILDKQYAEDSSNIKWIDAPPDCFPCFHWNNAWEEPEPDEVVYIGSCMTPPDIF 438
 Db 263 DLSKV-EGVIGLESQIKENTKDILESAVNPFRVRAKSKLGLQTESYYR-----TERNVDE- 362
 DR 263 DLSKV-EGVIGLESQIKENTKDILESAVNPFRVRAKSKLGLQTESYYR-----TERNVDE- 362

QY 439 N---ES-DENLKSYLSEI-----RLNIKTGSTERRPLISNEQQVLEA 478
 Db 310 GVIGGLESGIKENTKDILESAVNPFRVRAKSKLGLQTESYYR-----TERNVDE- 362
 DR 310 GVIGGLESGIKENTKDILESAVNPFRVRAKSKLGLQTESYYR-----TERNVDE- 362

QY 479 GMVYRNMLRKTKFA-YIYALAEPMVKVSGFARYKDLTTEEVKHLYGDNYRGGPFLEF-- 535
 Db 363 -----RVDRADQYAVYLILKHAGGKV-----KVKVQDYRE KYKPKVFLFQG 404
 DR 363 -----RVDRADQYAVYLILKHAGGKV-----KVKVQDYRE KYKPKVFLFQG 404

QY 536 -----GEGEDEEGYIYLCEVEDKTKWSELQIVNAVSEAVEATVKLSSRVPVGFHGTF- 588
 Db 405 KYIYAGESTYKNEE-----VKEIDALEIPNEI MRCGYEVIVPS-----HREFD 448
 DR 405 KYIYAGESTYKNEE-----VKEIDALEIPNEI MRCGYEVIVPS-----HREFD 448

QY 589 IGAD-DLAQOVV 599
 Db 449 IQRDVVDLIERIM 460
 DR 449 IQRDVVDLIERIM 460

RESULT 10

CYAA_USTMA STANDARD; PRT; 2493 AA.
 AC P49605;
 ID CYAA_USTMA
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl
 DE cyclase).
 DE UAC1 OR REM1.
 GN OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae;
 OC NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=518;
 RX MEDLINE=95087882; PubMed=7995519;
 RA Gold S., Duncan G., Barrett K., Kronstad J.W.;
 RT "cAMP regulates morphogenesis in the fungal pathogen Ustilago
 maydis";
 RL Genes Dev. 8:2805-2816 (1994).
 CC FUNCTION: PLAYS ESSENTIAL ROLES IN REGULATION OF CELLULAR
 CC METABOLISM BY CATALYSING THE SYNTHESIS OF A SECOND MESSENGER,
 CC CAMP.
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -!- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-3 FAMILY.

CC	- - SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).	Db	960 NTTNNATGTLPRNHE----IRVYKTDGTATLSCSLVSTANEVQ-----
CC	- - SIMILARITY: CONTAINS 1 PP2C-LIKE DOMAIN.	Db	998 NTNNATGTLPRNHE----IRVYKTDGTATLSCSLVSTANEVQ-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/) or send an email to licensee@isb-sib.ch).	Qy	217 DUGHTGTLRMLFYARAAGTIVDPAAHGTGVANGLWENGRLLAMSEDDLPYQVQ-----271
CC	DR InterPro; IPR01054; LRR.	Db	999 ---TIIARKSLTTEASYRLF----VRDG----SERPGIS---DKPSOLQRRLI 1041
CC	DR InterPro; IPR01611; LRR.	Qy	272 ---ITPNCNDLKTGYRFDFGQLESTMIAHPKVDPESGEFLAISYDVKPLKYFRSPD 328
CC	DR InterPro; IPR01592; LRR_out.	Db	1042 QAGTYTNDGLEMGRRDLS-----YLRLFVFRPD 1070
CC	DR InterPro; IPR01591; LRR_TYP.	Qy	329 GTKSPDVE-TQLDQPMMHDAITENVVVP 358
CC	DR InterPro; IPR01932; PP2C-like.	Db	1071 SVPTFESEISGHSEHTFQHLDLHSRNLEMVP 1101
CC	DR Pfam; PF00481; guanylate_cyc; 1.		
CC	DR Pfam; PF00560; LRR; 17.		
CC	DR SMART; SM0004; CYCC; 1.		
CC	DR SMART; SM00370; LRR; 6.		
CC	DR SMART; SM00369; LRR_TYP; 2.		
CC	DR SMART; SM00334; PP2C_C; 1.		
CC	DR SMART; SM00314; RA; 1.		
CC	KW PROSTE; PS50121; GUANYLATE_CYCLASES_2; 1.		
CC	KW Lyase; Repeat; Leucine-rich repeat; CAMP synthesis; Magnesium.		
FT	DOMAIN 759 763 POLY-ASP.	OS Brassica napus (Rape).	
FT	DOMAIN 888 897 POLY-ALA.	OC Viridiplantae; Streptophytta; Embryophyta; Tracheophyta;	
FT	REPEAT 1084 1106 LRR 1.	OC Spermatophytta; Magnoliophyta; eudicots; Rosidae;	
FT	REPEAT 1108 1131 LRR 2.	OC eurosids II; Brassicales; Brassicaceae; Brassica.	
FT	REPEAT 1132 1155 LRR 3.	OX NCBI_TAXID=3708;	
FT	REPEAT 1157 1178 LRR 4.	RN [1]	
FT	REPEAT 1179 1201 LRR 5.	RP SEQUENCE FROM N.A.	
FT	REPEAT 1202 1225 LRR 6.	RC STRAIN=cv; Drakkar; TISSUE=Seed;	
FT	REPEAT 1227 1248 LRR 7.	RA Zawodny S.; Martini N.;	
FT	REPEAT 1249 1271 LRR 8.	RT "Isolation and analysis of a cDNA clone encoding the small subunit of oilseed rape (Brassica napus)."	
FT	REPEAT 1273 1295 LRR 9.	RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.	
FT	REPEAT 1314 1337 LRR 10.	CC FUNCTION: THIS PROTEIN PLAYS A ROLE IN SYNTHESIS OF STARCH.	
FT	REPEAT 1338 1360 LRR 11.	CC IT CATALYZES THE SYNTHESIS OF THE ACTIVATED GLYCOSYL DONOR,	
FT	REPEAT 1361 1384 LRR 12.	CC ADP-GLUCOSE FROM GLC1-P AND ATP.	
FT	REPEAT 1386 1407 LRR 13.	CC -!- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate = diphosphate + ADP-glucose.	
FT	REPEAT 1408 1430 LRR 14.	CC -!- ENZYME REGULATION: ACTIVATED BY 3'-PHOSPHOGLYCERATE, INHIBITED BY ORTHOPHOSPHATE, ALLOSTERIC REGULATION.	
FT	REPEAT 1432 1455 LRR 15.	CC -!- PATHWAY: Starch biosynthesis.	
FT	REPEAT 1509 1529 LRR 16.	CC -!- SUBCELLULAR LOCATION: HETEROOTETRAMER.	
FT	REPEAT 1533 1556 LRR 17.	CC -!- TISSUE SPECIFICITY: LEAVES.	
FT	REPEAT 1557 1580 LRR 18.	CC -!- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-PHOSPHATE ADENYLYLTRANSFERASE FAMILY.	
FT	REPEAT 1581 1604 LRR 19.	CC	
FT	REPEAT 1606 1628 LRR 20.	CC	
FT	REPEAT 1633 1656 LRR 21.	CC	
FT	DOMAIN 1722 2001 PP2C-LIKE.	CC	
FT	SEQUENCE 2002 2493 AA; 271979 MW; 106A872C3C1CBF8 CRC64; CATALYTIC.	CC	
SQ		CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/).	
CC	Query Match 3.2%; Score 102; DB 1; Length 2493;	CC DR EMBL; AJ271162; CAB89863.1; -	
CC	Best Local Similarity 21.7%; Pred. No. 48; Mismatches 136; Indels 126; Gaps 17;	CC DR Pfam; PF00483; NTP_transf erase; 1.	
CC	Matches 85; Conservative 44; Mismatches 136;	CC DR PROSITE; PS00808; ADP_GLC_PYROPHOSPH_1; 1.	
CC	Qy 4 FTATAAVSGRLLGNHT ---QPLSSQSSQSDLSYCSSLIPASRYTRKLNVSSALHTP 57	CC DR PROSITE; PS00809; ADP_GLC_PYROPHOSPH_2; 1.	
CC	Db 801 FKTASLAGSRROTDDSYDPLTALPPLPGSKSYDEAAAKVDVLQQ TNNLQSQALVQQ- 858	CC DR PROSITE; PS00810; ADP_GLC_PYROPHOSPH_3; 1.	
CC	Qy 58 PALHFKKSSNSPAIVVPKAKESNTKOMNLFORAAAALDAEGELVSH ---EKHLPL 113	CC KW Glycogen biosynthesis; Transf erase; Nucleotidyltransferase;	
CC	Db 859 ---QSOSQNHQDPSNPVNPRTSPRGAGHMH - PASAGSAAAAGKUGLHRPSKHMNR 913	CC KW Chloroplast; Transit peptide; starch biosynthesis; Allosteric enzyme;	
CC	Qy 114 PKTAD -----PVQIAGNEAPVNEQPYVRNLPPVGKLPDSIKGVYVRNGANPLHPEVTG 167	CC FT TRANSIT 1 71 CHLOROPLAST (POTENTIAL).	
CC	Db 914 PNFGASVGPATRPTTGLGSTAEDD -----TSING1RDG_HPLKRSATA 959	CC FT CHAIN 72 520 GLUCOSE-1-PHOSPHATE ADENYLYLTRANSFERASE	
CC	Qy 168 -----BHFEDDGMMWHAVKFHGSSAYACRETQINRFVQERQLGRPEVKAGE 216		

OS	vicia faba	(Broad bean).
OC	Eukaryota	Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatozoids	Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurodids I	Papilionoideae; Fabaceae; Fabales; Fabaceae; Papilionoideae; Vicieae; Vicia.
OX	NCBI_TaxID=3906;	
[1]	RN	
	SEQUENCE FROM N.A.	
RC	RP	SEQUENCE FROM N.A.
RC	RC	SEQUENCE FROM COTYLEDON;
RX	RD	STRAIN=CV_Fribio; TISSUE=COTYLEDON;
RA	RD	MEDLINE=952188601; PubMed=7766042;
RA	RA	Weber H., Heim U., Borisjuk L., Wobus U.;
RT	RA	Cell-type specific, coordinate expression of two ADP-glucose
RT	RT	pyrophosphorylase genes in relation to starch biosynthesis during
RT	RT	seed development of <i>Vicia faba</i> L.";
RL	RT	<i>Planta</i> 195:352-361(1995).
RL	CC	-/- FUNCTION: THIS PROTEIN PLAYS A ROLE IN SYNTHESIS OF STARCH.
CC	CC	IT CATALYSES THE SYNTHESIS OF THE ACTIVATED GLYCOSYL DONOR,
CC	CC	ADP-GLUCOSE FROM GLC-1-P AND ATP.
CC	CC	-/- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
CC	CC	diphosphate + ADP glucose.
CC	CC	-/- ENZYME REGULATION: ACTIVATED BY 3'-PHOSPHOGLYCERATE, INHIBITED
CC	CC	BY ORTHOPHOSPHATE ALLOSTERIC REGULATION.
CC	CC	-/- PATHWAY: Starch biosynthesis.
CC	CC	-/- SUBCELLULAR LOCATION: HETEROPLASM.
CC	CC	-/- SUBCELLULAR LOCATION: Chloroplast. (By similarity).
CC	CC	-/- TISSUE SPECIFICITY: LEAVES AND SEEDS.
CC	CC	-/- DEVELOPMENTAL STAGE: IT IS PRESENT IN YOUNG COYLEDONS AT 14 DAYS
CC	CC	AFTER FERTILIZATION (DAF) WHEN CELLS ARE STILL RAPIDLY DIVIDING.
CC	CC	LEVELS STEADILY ACCUMULATE UNTIL THE END OF THE CELL EXPANSION
CC	CC	PHASE (35-40 DAF), AND WITH THE BEGINNING OF THE SEEDS DEPICTICATION
CC	CC	PHASE (AT 50 DAF), THE LEVELS DECREASE TO VERY LOW LEVELS.
CC	CC	-/- SIMILARITY: BELONGS TO THE BACTERIAL AND PLANTS GLUCOSE-1-
CC	CC	PHOSPHATE ADENYLTRANSFERASE FAMILY.
CC	CC	-----
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	CC	the European Bioinformatics Institute. There are no restrictions on
CC	CC	use by non-profit institutions as long as its content is in no
CC	CC	modified and this statement is not removed. Usage by and for commercial
CC	CC	entities requires a license agreement (see http://www.isb-sib.ch/annex)
CC	CC	or send an email to license@isb-sib.ch).
CC	CC	-----

DR	EMBL-X76941; CAA54260.1; -.			
DR	Integ_Pro; IPR01825; NTP transferase;			
DR	Pfam:PF00483; NTP transferase; 1.			
DR	PROSTATE; PS00088; ADP_GLC_PYROPHOSPH_1; 1.			
DR	PROSTATE; PS00089; ADP_GLC_PYROPHOSPH_2; 1.			
DR	PROSTATE; PS00080; ADP_GLC_PYROPHOSPH_3; 1.			
KW	Glycogen biosynthesis; Transferase; Nucleotidyltransferase;			
KW	Multigene family; starch biosynthesis; Allosteric enzyme;			
KW	Chloroplast; starch biosynthesis; Allosteric enzyme;			
TRANSIT	1	63	CHLOROPLAST (POTENTIAL).	
CHAIN	64	512	GLUCOSE-1-PHOSPHATE ADENYLYLTRANSFERASE	
			SMALL SUBUNIT 2.	
SEQUENCE	512 AA:	56059 MW:	060547F683526EDC CRC64:	
FT				
SQ				
Query Match	3 / 2%	Score 101; DB 1; Length 512;		
Best Local Similarity	21.0 %	Pred. No. 5.7;		
Matches	118;	Conservative 68; Mismatches 170; Index 206; Caps		
30;				
QY	22 PPLSSQSQDSLXCSSSLPMASRVRKLVNSALHTPPALHF-----PKQSSNSPAIVVKP	76		
Db	10 PPSSSSSSSS-----SKAARNLNLSSS-HLSGDKFTLGSRTRRSGRNPFIVSP	61		
Qy	77 KA-KESINTSKQMNLEQRAAAALDAEGFLV--SHEKLHL-PKTADEPSVQIAGNEFAPVN	131		
Db	62 KAVSDSKNSQTCLPDDASRSLV---GILGGGACTRLYPLTKRAKPAPVLGANYRLI-	116		
QY	132 EQPVRRNLPVVGKLPDSIKGVYVRNGANPLHEPYTGHHF-----PGDGHIVHAKV	182		
Db	117 -----DIPVSNCNSNTSKIYVLTQFN--SASLNRLHSRAYASNLGGYKNEGVEVLA	167		
QY	182 FE-----HGSA----SYACRETQTN-----REVOEROTGRPFPEKATG	215		

EMBR: X54853 CAA39621

		3-phytate precursor (EC 3.1.3.8) (Phytate 3'-phosphatase) (MYO-DE
		inosititol-hexaphosphate 3'-phosphotyrosinase)
		GN PHYC OR PHYB13.
		bacillus subtilis.
		Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus. 4
		NCBI_TAXID=1423;
RN		SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP		STRAIN=NATV-E-68013;
RC		MEDLINE 98268943; PubMed=9603817;
RX		Kerouvo J., Lauriainen M., Nurminen P., Kalkkinen N., Apajaalahti J.;
CC		"Isolation, characterization, molecular gene cloning, and sequencing of a novel phytase from <i>Bacillus subtilis</i> ."
CC		Appl. Environ. Microbiol. 64:2079-2085(1998).
CC		-!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHYTATE. ONLY PHYTATE, ADP, AND ATP WERE HYDROLYZED (100, 75, AND 50% OF THE RELATIVE ACTIVITY, RESPECTIVELY). HAS MAXIMAL ACTIVITY AT PH 7 AND 55 DEGREES CELSIUS.
CC		-!- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = D-myoinositol 1,2,4,5,6-pentakisphosphate + phosphate.
CC		-!- COFACTOR: ISOLATED ENZYME REQUIRED CALCIUM FOR ITS ACTIVITY AND/OR STABILITY
CC		-!- SUBCELLULAR LOCATION: Secreted.
CC		-!- INDUCTION: BY PHYTATE.
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@lsb-sib.ch).
CC		EMBL: AF029053; AAC31775.1; -.
DR	HSSP; IPR004343; Phytase.	DR HSSP; 066037; 2POO.
DR	InterPro; IPR00379; Ser-esterase.	InterPro; IPR00379; 2POO.
DR	InterPro; IPR01031; Thioesterase.	InterPro; IPR01031; 2POO.
DR	InterPro; IPR01242; Condensatn.	InterPro; IPR01242; 2POO.
DR	InterPro; IPR003880; Phntate-attach.	InterPro; IPR003880; 2POO.
DR	InterPro; IPR000501; AMP-binding.	InterPro; IPR000501; 2POO.
DR	PROTINSI; PRO00154; AMPBINDING.	PROTINSI; PRO00154; 2POO.
DR	PROSITE; PS00012; PHOSPHOPANTETHINE; 3.	PROSITE; PS00012; 2POO.
DR	PROSITE; PS00455; AMP-BINDING; 3.	PROSITE; PS00455; 2POO.
DR	PROSITE; PS50075; ACP-DOMAIN; 3.	PROSITE; PS50075; 2POO.
KW	Ligase; Antibiotic biosynthesis; Multifunctional enzyme;	
KW	REPEAT; Phosphopantethine.	
REPEAT	321 910 DOMAIN 1 (ADIPATE-ACTIVATING).	
REPEAT	1413 1993 DOMAIN 2 (CYSTEINE-ACTIVATING).	
REPEAT	2494 3078 DOMAIN 3 (VALINE-ACTIVATING).	
DOMAIN	850 919 ACYL CARRIER (ACP) 1.	
DOMAIN	2020 2002 ACYL CARRIER (ACP) 2.	
DOMAIN	3020 3087 ACYL CARRIER (ACP) 3.	
BINDING	882 882 PHOSPHOPANTETHINE (BY SIMILARITY).	
BINDING	1965 1965 PHOSPHOPANTETHINE (BY SIMILARITY).	
FET	3050 3050 PHOSPHOPANTETHINE (BY SIMILARITY).	
FET	3623 3623 THIOESTERASE (BY SIMILARITY).	
ACT_SITE	SEQUENCE 3770 AA; 422448 MW; CB66B6D232A58CB0 CRC64;	
Query Match Score 3.2%; Best Local Similarity 20.9%;保守性 97%; Matches 67; Gaps 27; Query 12 GRWLGHNHTQQPLSSQSSDLSYCSLPMASRVRTRK-----INVSALHIPPAL 60	61 -----HEPKQSS---NSPAIIVKPKAKESNTYQMNLFQRAAAALDA 99	61 -----HEPKQSS---NSPAIIVKPKAKESNTYQMNLFQRAAAALDA 99
QY	1544 GNWSVSN---PAPSSTDLAY--IIYTSGTGKPKGVAVFHGVNLQISLSKTFGL 1596	1544 GNWSVSN---PAPSSTDLAY--IIYTSGTGKPKGVAVFHGVNLQISLSKTFGL 1596
Db	1597 RDTDDDEVLSFSNVFDFHEVQMDAILNGOTLWLDAMRSRKR--LYQYIENNRVTY 1654	1597 RDTDDDEVLSFSNVFDFHEVQMDAILNGOTLWLDAMRSRKR--LYQYIENNRVTY 1654
QY	100 AEGFLVSHEKHLPLPKTADPSVQIAGNFAPVNQFVR-----NLPVGKLELDSIKG 151	100 AEGFLVSHEKHLPLPKTADPSVQIAGNFAPVNQFVR-----NLPVGKLELDSIKG 151
Db	1655 LSG-----TPSIVMSYEFSRFKDH LRRYDCVGEEAFSQVFDQDQTDFOQ 1698	1655 LSG-----TPSIVMSYEFSRFKDH LRRYDCVGEEAFSQVFDQDQTDFOQ 1698
QY	152 VYVRGANPLHEPVGHFFPDGDMMHAVKEFHSSAYACRFQNRREVQRLGR-PVF 210	152 VYVRGANPLHEPVGHFFPDGDMMHAVKEFHSSAYACRFQNRREVQRLGR-PVF 210
Db	1699 LII NGYSPTEISITHK-----RYPPEPRRDKSIGQIGNSTSYVLNADMRRVPI- 1750	1699 LII NGYSPTEISITHK-----RYPPEPRRDKSIGQIGNSTSYVLNADMRRVPI- 1750
QY	211 PKAIGELH-GHTGTLARMLFYAR---AAGIVDPAHGTYVANAGLVYENGRLLAMSDD 265	211 PKAIGELH-GHTGTLARMLFYAR---AAGIVDPAHGTYVANAGLVYENGRLLAMSDD 265
Db	1751 -GAVGELYLGGEGYAR-GYHNRPEVTAERFLRNQFQTDSERONGR--NSRLYRTGD- 1802	1751 -GAVGELYLGGEGYAR-GYHNRPEVTAERFLRNQFQTDSERONGR--NSRLYRTGD- 1802
QY	266 LPYQQTIP-NGDITKYGFR-----GOLESMTIAH-----KVDPES 304	266 LPYQQTIP-NGDITKYGFR-----GOLESMTIAH-----KVDPES 304
Db	1803 --LWRWPGNSEGEYLRNDFQWIKIRGLRIEIEAMSSHHDIKQSVVIAKGKEG 1859	1803 --LWRWPGNSEGEYLRNDFQWIKIRGLRIEIEAMSSHHDIKQSVVIAKGKEG 1859
QY	305 GELFALSDDVSKPYLKYFRESPDGTKSPDVETQLDQPTMMHDFAITENEV----VVP 358	305 GELFALSDDVSKPYLKYFRESPDGTKSPDVETQLDQPTMMHDFAITENEV----VVP 358
Db	1860 DQKFLVGKEVASSP-----LSPGA-----IRFMOSSLPGYMPMSSEPISSLPVTP 1906	1860 DQKFLVGKEVASSP-----LSPGA-----IRFMOSSLPGYMPMSSEPISSLPVTP 1906
QY	359 DQOVFFK-LPEM-TRGSPVYDKNKVARE----GILDKYAE 394	359 DQOVFFK-LPEM-TRGSPVYDKNKVARE----GILDKYAE 394
Db	1907 SGKLTAKALPTAERGMVNLAPRNIEIESILCGISAGLDISAQ 1950	1907 SGKLTAKALPTAERGMVNLAPRNIEIESILCGISAGLDISAQ 1950
RESULT 15		
PHYC_BACSU	STANDARD;	PRT; 383 AA.
ID	AC_031097; 30-MAY-2000 (Rel. 39, Created)	AC_031097; 30-MAY-2000 (Rel. 39, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last sequence update)	30-MAY-2000 (Rel. 39, Last sequence update)
DB	356 VWPQOVVYFKLPENRGSPVVY---KNG-VAREGILDKYAEDSSNNIKWIDAPCFCH 411	356 VWPQOVVYFKLPENRGSPVVY---KNG-VAREGILDKYAEDSSNNIKWIDAPCFCH 411

Tue May 20 09:35:03 2003

us-09-758-269-6.rsp

Page 13

Db 275 MASSQ-----GNSSYAIYDROGKNNVADFRITDGPETDGTS----- 311
Qy 412 LWNAWDEPEDEVVVIGSCMP--PDSTFINESD-ENL 445
Db 312 -----DTDGIDVLGFLGPEYFGIFVAQDGENI 340

Search completed: May 19, 2003, 15:48:01
Job time : 45 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 15:46:14 ; Search time 43 Seconds
 Perfect score: 3150 ; 1 MASFTATAAVSGRWLGGNHT.....VYGFHGFIGADDLAKQVY 599
 Sequence: Q8YXN9

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671596

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : SPTREMBL_21;*

- 1: sp_archaea;*
- 2: sp_bacteria;*
- 3: sp_fungi;*
- 4: sp_invertebrate;*
- 5: sp_mammal;*
- 6: sp_mitochondrion;*
- 7: sp_nucleus;*
- 8: sp_organelle;*
- 9: sp_phage;*
- 10: sp_plant;*
- 11: sp_rabbit;*
- 12: sp_virus;*
- 13: sp_vertebrate;*
- 14: sp_unclassified;*
- 15: sp_rvirus;*
- 16: sp_bacteriaph;*
- 17: sp_archaea;*

17 937 29.7 538 10 065572 arabidopsis
 18 837.5 26.6 524 10 Q9AEN8 lactuca sativa
 19 798.5 25.3 475 10 Q9AXZ5 perseae amer
 20 745.5 23.7 446 10 Q8XXV3 anaibaena sp
 21 745 23.7 446 10 Q49895 malus domestica
 22 528.5 16.8 483 16 Q9AA32 caulobacter
 23 527.5 16.7 456 2 Q9AM11 streptomyces
 24 526.5 16.7 456 2 Q93FA4 streptomyces
 25 466 14.8 501 16 Q06785 mycobacterium
 26 464 14.7 503 16 Q9RK46 streptomyces
 27 454.5 14.4 502 16 Q05905 mycobacterium
 28 394 12.5 490 16 P74334 synechocystis
 29 382.5 12.1 209 10 Q8YXN9 crocus sativus
 30 355.5 11.3 497 2 Q8YPB4 anabaena sp
 31 352 11.2 485 2 Q5353 pseudomonas
 32 327 10.4 218 10 Q8VXP1 crocus sativus
 33 322.5 10.2 570 10 Q8AY26 arabidopsis
 34 319 10.1 520 16 Q8TA2 ralstonia S
 35 317 10.1 480 16 P74370 synechocystis
 36 311.5 9.9 491 2 Q8CQW2 pseudomonas
 37 303.5 9.6 490 2 Q52008 pseudomonas
 38 302.5 9.6 552 10 Q93VDS oryza sativa
 39 270 8.6 342 2 Q68668 synechococcus
 40 260 8.3 526 13 Q91993 gallus gallus
 41 256 8.1 516 13 Q9NW44 brachydanois
 42 253 8.0 616 10 Q9M079 arabidopsis
 43 244 7.7 618 10 Q9SHD9 arabidopsis
 44 242 7.7 185 10 Q8VXP0 crocus sativus
 45 242 7.7 485 16 Q8YMN8 anabaena sp

ALIGNMENTS

RESULT 1
 Q9LRR7 PRELIMINARY; PRT; 599 AA.
 ID Q9LRR7; AC Q9LRR7; DT 01-OCT-2000 ("TREMBL"; 15, Created)
 AC Q9LRR7; DT 01-OCT-2000 ("TREMBL"; 15, Last sequence update)
 AC Q9LRR7; DT 01-DEC-2001 ("TREMBL"; 19, Last annotation update)
 DE 9-cis-epoxycarotenoid dioxygenase (Neoxanthin cleavage enzyme).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;
 OC Spermatophytina; Magnoliophytina; eudicots; Rosidae;
 OC Eurosids II; Brassicales; Brassicaceae; Arabidopsidae;
 OC NCBITaxonID=3702; OX NCBI_TaxID=3702; RN [1]
 RN SEQUENCE FROM N. A.
 RC STRAIN=COLUMBIA;
 RA Sato S.; Nakamura Y.; Kaneko T.; Asamizu E.; Tabata S.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N. A.
 RC SPRAIN-COL.; RA Tuchi S.; Kobayashi M.; Shinozaki K.;
 RX MEDLINE=20277480; PubMed=10819329; RA Nakamura Y.;
 RA "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 clones."; RT DR Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RL DR EMBL; AB028617; BAB01336; 1.
 RT DR InterPro; IP0004234; RPE65.
 RT DR Pfam; PF03055; RP65; 1.
 KW Dixygenase.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Query	Match	Length	DB	ID	Description
1	3150	100.0	599	10	Q9LRR7	Q9lrr7 arabidopsis
2	3143	99.8	599	10	Q93ZU5	Q93zus3 solanum tuberosum
3	2289	72.7	604	10	Q9M329	Q24023 lycoopersico
4	2280.5	72.4	605	10	Q9M9F5	Q9m9f5 arabidopsis
5	2170.5	68.9	657	10	Q9FS24	Q9fs24 vigna unguifera
6	2168.5	68.8	612	10	Q9M6E8	Q9m6e8 phassolus vulgaris
7	2165.5	68.7	615	10	Q9AXZ4	Q9axz4 perseae amer
8	2066.5	65.6	625	10	Q96ZG1	Q9c6z1 arabidopsis
9	2044.5	64.9	589	10	Q49505	Q49505 arabidopsis
10	1991	63.2	583	10	Q49105	Q9axz3 perseae amer
11	1977	62.8	569	10	Q24592	Q24592 zea mays (improved inb3)
12	1930	61.3	604	10	Q9IRM7	Q9irm7 arabidopsis
13	1663.5	52.8	577	10	Q49675	Q49675 arabidopsis
14	1668	30.7	595	10	Q9ATR2	Q9air2 phassolus vulgaris
15	959.5	30.5	543	10	Q9LY63	Q9ly63 arabidopsis
16	938	29.8	536	10		

SQ	SEQUENCE	599 AA;	65856 MW;	7D513E39945B0CF3 CRC64;	DR	EMBL; AY056255; AAL07104.1; -.
	Query Match	100.0%;	Score 3150;	DB 10;	DR	InterPro; IPR004294; RPE65.
	Best Local Similarity	100.0%;	Pred. No. 1	8e-237;	DR	Pfam; PF03055; RPE65; 1.
	Matches 599;	Conservative 0;	Mismatches 0;	Indels 0;	KW	Dioxygenase.
	Db	1	MASFTATAAVSGRNGLGGNHTOPPLSSQSSDSYCSSLPMASRVTRKLNSVSSALHTPPAL 6.0	SEQUENCE	599 AA;	65813 MW;
QY	1	MASFTATAAVSGRNGLGGNHTOPPLSSQSSDSYCSSLPMASRVTRKLNSVSSALHTPPAL 6.0	Query Match	99.8%;	Score 3143;	DB 10;
Db	1	MASFTATAAVSGRNGLGGNHTOPPLSSQSSDSYCSSLPMASRVTRKLNSVSSALHTPPAL 6.0	Best Local Similarity	99.8%;	Pred. No. 6.5e-237;	Length 599;
QY	61	HFPKQSSNSPATAVVKPKAKESNTKOMNLQRAAAAALDAEAEFLVSHESKEKLHLPKTAAPS 1.20	Matches 598;	Conservative 0;	Mismatches 1;	Indels 0;
Db	61	HFPKQSSNSPATAVVKPKAKESNTKOMNLQRAAAAALDAEAEFLVSHESKEKLHLPKTAAPS 1.20	SEQUENCE	599 AA;	65813 MW;	Gaps 0;
QY	121	VQAGNFAPVNQSPVRLNPVYKGKLPDSIKGTVVRNGANPLHEPVTCGHFFGDGMVHAV 1.80	Query Match	99.8%;	Score 3143;	DB 10;
Db	121	VQAGNFAPVNQSPVRLNPVYKGKLPDSIKGTVVRNGANPLHEPVTCGHFFGDGMVHAV 1.80	Best Local Similarity	99.8%;	Pred. No. 6.5e-237;	Length 599;
QY	181	KFEGGSASYACRFQTNRFQRQLGRPVPEPAIGEDEHGHTGTRALMLFYARAAGTVDP 2.40	Matches 598;	Conservative 0;	Mismatches 1;	Indels 0;
Db	181	KFEGGSASYACRFQTNRFQRQLGRPVPEPAIGEDEHGHTGTRALMLFYARAAGTVDP 2.40	SEQUENCE	599 AA;	65813 MW;	Gaps 0;
QY	241	AHGTGVANAGLYVFNGRLLAMSEDDLFPLYQVOTTPNGDLKTYGRDFDQESTSMIAHPKV 3.00	Query Match	99.8%;	Score 3143;	DB 10;
Db	241	AHGTGVANAGLYVFNGRLLAMSEDDLFPLYQVOTTPNGDLKTYGRDFDQESTSMIAHPKV 3.00	Best Local Similarity	99.8%;	Pred. No. 6.5e-237;	Length 599;
QY	301	DEPSGEFALSYYDSVSKPLYKFRESPDGTKSPDVETQDQPTMMHDEAITEENFVVVPDQ 3.60	Matches 598;	Conservative 0;	Mismatches 1;	Indels 0;
Db	301	DEPSGEFALSYYDSVSKPLYKFRESPDGTKSPDVETQDQPTMMHDEAITEENFVVVPDQ 3.60	SEQUENCE	599 AA;	65813 MW;	Gaps 0;
QY	361	QVVFKLPPEMIRGSPVYYDKKKVAREGILDKYAEDESSNIKWDADPDCFHWWNAMEEPE 4.20	Query Match	99.8%;	Score 3143;	DB 10;
Db	361	QVVFKLPPEMIRGSPVYYDKKKVAREGILDKYAEDESSNIKWDADPDCFHWWNAMEEPE 4.20	Best Local Similarity	99.8%;	Pred. No. 6.5e-237;	Length 599;
QY	421	TDEVVVIGSCMTPPDSTFNESDENLKSVLSETRLNLTGESTTRPITNSEDQVNLEAGM 4.80	Matches 598;	Conservative 0;	Mismatches 1;	Indels 0;
Db	421	TDEVVVIGSCMTPPDSTFNESDENLKSVLSETRLNLTGESTTRPITNSEDQVNLEAGM 4.80	SEQUENCE	599 AA;	65813 MW;	Gaps 0;
QY	481	VNRNMGRKTKEYAYLAEPWKVSGFAKVDLTTGEVKHHLYGDNYGGEEFLFEGEGE 5.40	Query Match	99.8%;	Score 3143;	DB 10;
Db	481	VNRNMGRKTKEYAYLAEPWKVSGFAKVDLTTGEVKHHLYGDNYGGEEFLFEGEGE 5.40	Best Local Similarity	99.8%;	Pred. No. 6.5e-237;	Length 599;
QY	541	EDEGYLICFVHDEKTWKSELQIVNAVSLEYEATVKLPSRVYPGFHGTFIGADDLAKQVV 5.99	Matches 598;	Conservative 0;	Mismatches 1;	Indels 0;
Db	541	EDEGYLICFVHDEKTWKSELQIVNAVSLEYEATVKLPSRVYPGFHGTFIGADDLAKQVV 5.99	SEQUENCE	599 AA;	65813 MW;	Gaps 0;
RESULT 2					RESULT 3	
Q93ZU5	ID	PRELIMINARY;	PRT;	599 AA.	Q9M3Z9	PRELIMINARY;
AC	Q93ZU5				ID	PRT;
DT	01-DEC-2001	(TREMBLrel. 19, Created)			Q9M3Z9	604 AA.
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			AC	
DE	01-MAR-2002	(TREMBLrel. 20, Last annotation update)			DT	01-OCT-2000 (TREMBLrel. 15, Created)
GN	MOA2 ./AR3G14440.	Putative 9-cis-epoxycarotenoid dioxygenase.			DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
OS	Arabidopsis thaliana (Mouse-ear cress).				DB	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				GN	Putative 9-cis-epoxycarotenoid dioxygenase.
OC	eu dicots; Magnoliophyta; eudicots; core eudicots;				NCBI_TAXID	NCED1.
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.				OS	Solanum tuberosum (Potato).
RN	[1]	SEQUENCE FROM N.A.			OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
RA	Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,				Sp	Asteridae; eu dicots I; Solanales; Solanaceae; Solanum.
RA	Onderka C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,				NCBI_TAXID	4113;
RA	Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,				OX	
RA	Ishida J., Jones T., Kamaya A., Karlin-Neumann G., Kawai J., Kim C.,				RN	
RA	Koescen E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,				RP	
RA	Nguyen M., Palm C.J., Sakurai T., Satou M., Shinnin P.,				RA	SEQUENCE FROM N.A.
RA	Southwick A., Tracy S.E., Shiozaki K., Davis R.W., Ecker J.R.,				RA	Burbridge A.; Taylor I.B.; Thompson A.;
RA	Theologis A.;				RT	"Potato putative 9-cis-epoxycarotenoid dioxygenase 1 cDNA."
RT	"Full Length cDNA of gene MOA2.4/At3g14440 (GI:11994214)." ;				RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.				DR	InterPro; IPR004294; CAB6920.1; -.
					DR	PFam; PF03055; RPE65; 1.

KW	Dioxygenase.	RP	SEQUENCE FROM N.A.
SQ	SEQUENCE 604 AA; 67288 MW; 36856BC82E160A1 CRC64;	RA	Burbridge A.;
	Query Match Score 2289; DB 10; Length 604;	RL	Submitted to the EMBL/GenBank/DBJ databases.
	Best Local Similarity 71.3%; Pred. No. 3.5e-170;	DR	EMBL; Z97215; CAB10168.1;
Matches 432; Conservative 66; Mismatches 90; Indels 18; Gaps 5;	DR	InterPro; IPR004294; RPE65.	
	DR	Pfam; PF03055; RPE65; 1.	
	KW	Dioxygenase.	
QY	5 TATAAVSGRGLGGNHTOPPLSSQSSDLSYCS--SLPMASRVKLNYSALHTPPALH 61	SEQUENCE	605 AA; 67317 MW; 19BA67930346A872 CRC64;
Db	3 TTTTAAATTTW----KPKLSSPSKEFASNSISLKKNNHQSLNNSLQQAPLH 57	Query Match	Score 2280.5; DB 10; Length 605;
QY	62 EPKQSN---SPAIVVKPKAKESNT---KQMLNLFQRAAAALDAAEGLFVSHKEKLHPL 113	Best Local Similarity	72.4%; Pred. No. 1.6e-169;
Db	58 EPKQSSNYOTPKTSTISHPKTNSSSSISKWNLVQKAAMALDAVEGALTKEELHPL 117	Matches	435; Conservative 63; Mismatches 90; Indels 19; Gaps 6;
QY	114 PKTADPSPVQLAGNAPVNQPVRLNPVQKLPSSIKGYVRNQANPHEPTVGHHTFDG 173	QY	5 TATAAVSGRGLGGNHTOPPLSSQSSDLSYCS--SLPMASRVKLNYSALHTPPALH 61
Db	118 PKTADPVPQISGNAPVPPNPVCOSLPVTGKIPKCQGVYVRNQANPHEPTVGRHHTFDG 177	Db	3 TTTSIAATNTW----KPKLSSPSKEFASNSISLKKNNHQSLNNSLQQAPLH 57
QY	174 DGMVHAVKEFPGSAYASACRFTOTNRREVQERQLGRPVFPEPKAIGRLHGTGIAARLMLEYARA 233	QY	62 EPKQSN---SPA--IVVKPKAKESN---TKQMNLFQRAAAALDAAEGLFVSHKEKLHPL 112
Db	178 DGMVHAVKEFPGSAYASACRFTOTNRREVQERQLGRPVFPEPKAIGRLHGTGIAARLMLEYARG 237	Db	58 EPKQSSNYOTPKTNSSSSISKWNLVQKAAMALDAVEGALTKEELHPL 117
QY	234 AAGIVDPAHGTGVAAGLVYENGRLLANSEDDLPYQVQITPNDLKVTFGRFDGQLEST 293	QY	113 LPKTADPSPVQLAGNAPVNQPVRLNPVQKLPSSIKGYVRNQANPHEPTVGHHTFDG 172
Db	238 LFLGDIDHRSRGTVGAGLVYENRNLAMSEDDLPYHVKVTPSDLKTEGREFEDGQLKST 297	Db	118 LPKTADPVPQISGNAPVPPNPVCOSLPVTGKIPKCQGVYVRNQANPHEPTVGRHHTFDG 177
QY	294 MIAHPKVDPEGLEALSDVSKPLKTFKFRRFSPDGKSPDQIQLDQMMDFATEEN 353	QY	173 DGMVHAVKEFPGSAYASACRFTOTNRREVQERQLGRPVFPEPKAIGRLHGTGIAARLMLEYAR 232
Db	298 MIAHPKLDPPVGSELFALSYDVKQPKLYFRFSKGERSNSDVEIPDPTMMHDEATEK 357	Db	178 GDGMVHAVQFKNGASYARFTETEVLQEKALGRPVFPEPKAIGRLHGTGIAARLMLEYAR 237
QY	354 FVYVPDQOWFKLPEMTRGSPPVYDKNVAEFLGKSPVYDNKVSFPGTILKQYKAKGSDLAWEVPCFCPHLW 413	QY	233 AAAGIVDPAHGTGVAAGLVYENGRLLANSEDDLPYQVQITPNDLKVTFGRFDGQLES 292
Db	358 FVYVPDQOWFKMSEMLFGSPVYDNKVSFPGTILKQYKAKGSDLAWEVPCFCPHLW 417	Db	238 GLFGLVLDHSKGTVGAGLVYENRNLAMSEDDLPYHVKVTPSDLKTEGREFEDGQLKST 297
QY	414 NAMEEPETDEVVWIGSCMTPPDSTFNESENLENKVLSEIRLNKGTGESTRPPIISNEDQ 473	QY	293 TMAHPKVDPESEGLEALSDVSKPLKTFKFRRFSPDGKSPDQIQLDQMMDFATEE 352
Db	418 NAMEEPETDEVVWIGSCMTPPDSTFNESENLENKVLSEIRLNKGTGESTRPPIISNEDQ 477	Db	298 TMAHPKLDPPVGSELFALSYDVKQPKLYFRFSKGERSNSDVEIPDPTMMHDEATE 357
QY	474 VNLLEAGMVNRNMUGRKTKFAYALAAEPWKVSGFAKDLTQEUVKLYGDNRYGGPLF 533	QY	353 NFVYVPDQOVVFKLPEMTRGSPPVYDKNVAEFLGKSPVYDNKVSFPGTILKQYKAKGSDLAWEVPCFCPHL 412
Db	478 VNLLEAGMVNRNMUGRKTKFAYALAAEPWKVSGFAKDLTQEUVKLYGDNRYGGPLF 537	Db	358 NFVYVPDQOVVFKLPEMTRGSPPVYDKNVAEFLGKSPVYDNKVSFPGTILKQYKAKGSDLAWEVPCFCPHL 417
QY	534 LPGE--AGEEDRGTYLQFHDEKTWKSELQTVNAVSLVEAVKLPSPRVYGFGRGIFTGA 591	QY	413 WNAWEEPBTDEVVWIGSCMTPPDSTFNESENLENKVLSEIRLNKGTGESTRPPIISNEDQ 472
Db	538 LPRDPSNEKEEDGYLQAFVHDEKEWSEIQLVNAVSLVEAVKLPSPRVYGFGRGIFTGA 597	Db	418 WNAWEABTDEVVWIGSCMTPPDSTFNESENLENKVLSEIRLNKGTGESTRPPIISNEDQ 477
QY	592 DDLAKQ 597	QY	473 QNLNEAGMVNRNMUGRKTKFAYALAAEPWKVSGFAKDLTQEUVKLYGDNRYGGPL 532
Db	598 NDLANQ 603	Db	478 QNLNEAGMVNRNMUGRKTKFAYALAAEPWKVSGFAKDLTQEUVKLYGDNRYGGPL 537
RESULT 4			
O24023	PRELIMINARY;	PRT;	605 AA.
AC		QY	591 ADDLAKQ 597
DT	01-JAN-1998 (TRIMBLrel. 05, Created)	ID	Q9M9F5
DT	01-JUN-1998 (TRIMBLrel. 06, Last sequence update)	AC	Q9M9F5;
DT	01-DEC-2001 (TRIMBLrel. 19, Last annotation update)	DT	01-OCT-2000 (TRIMBLrel. 15, Created)
DS	Nine-cis-epoxycurcetinoid dioxygenase.	DT	01-OCT-2000 (TRIMBLrel. 15, Last sequence update)
OC	Lycopersicon esculentum (Tomato).	DE	F3F9_10.
OC	Spermatophyta; Magnoliophyta; Streptophyta; Embryophyta; Tracheophyta;	OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Asteridae; euasterids I; Solanaceae; core eudicots;	OC	Spermatophyta; Magnoliophyta; eudicots; Rosidae;
OX	NCBI_TAXID=4081;	OX	eurosid II; Brassicales; Brassicaceae; Arabidopsis.
RN	SEQUENCE FROM N.A.	RN	[1] NCBI_TAXID=3702;
RN	"Structure and expression of a cDNA encoding a putative neoxanthin	RN	[2] J. Exp. Bot. 47:2111-2112(1997).
RN	cleavage enzyme (NCE) isolated from a wild-related tomato "	RN	
RN	(Lycoopersicon esculentum Mill.) library.";	RN	
RN	J. Exp. Bot. 47:2111-2112(1997).	RN	
RN	SEQUENCE FROM N.A.	RN	

Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C., Shinn P., Altaifi H., Bei P., Chin C., Chiu J., Choi E., Conn L., Conway A., Gonzales A., Hansen N., Hong B., Koc T., Lam F., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Toriumi M., Vaysberg M., Yu G., Federpiel N.A., Federpiel N.A., Ecker J.R.; "Genomic sequence for Arabidopsis thaliana BAC F3F9 from chromosome I." [2]	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN	SEQUENCE FROM N.A.
RN	Ecker J.R.; Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[3]	SEQUENCE FROM N.A.
Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C., Altaifi H., Bei P., Chin C., Chiu J., Choi E., Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam F., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Theologis A., Ecker J.; Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AC013430; AAC017197; "-"
DR	InterPro; IPR001899; Gram_pos_anchor.
DR	InterPro; IPR004294; RPE65.
DR	Pfam: PF01055; RPE65_1.
DR	PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ	SEQUENCE: 657 AA; 73015 MW; F41DCEBE94806318 CRC64;
Query	Query Match 68.9%; Score 2170.5; DB 10; Length 657;
Db	Best Local Similarity 67.7%; Pred. No. 7.2e-161;
Matches 419; Conservative 80; Mismatches 87; Indels 33; Gaps	
Qy	1 MASFTATAA YSGRWL ----- GGNHTQPPLSSQS QDSLVS CSCLP----- MASRYT 4;
Db	51 MASTULLPSQTSDTRTE STSSSSRSPKLSQ----- LSSSTSRNKKLVPCYVSSSN 1
Qy	46 RKLVYSSALH----PPALHPPQSSNSPATIVVKPKAKESNTKQMNLQRQAALDAAE 1
Db	106 KSSVSSSLLSQSPTEKPSW---KKLCNDVTLN1-PKTTNQNPK-LNPVORTAMYLDAVE 1
Qy	102 GFTVSH-EKHLPLKPTADPSYQIAGNFPAVNPQYVRLPNLPVGGKLDSTIKGVYVNGANP 1
Db	161 NMISHERRRRHHPKTDAPVQIAGNFPAVNPQYVRLPNLPVGTPECIGVYVNGANP 2
Qy	161 LHEPYTGHHFFGDGMAYHAYKFEGHSASYACRFETONRVEOQLGRPEFQAGLHLGH 2
Db	221 LKPVSCHHLEFGDMAYHAYRFDGSYVSYACFETENRVLRECRCPVFKAIGLHLGH 2
Qy	221 TGTARMLFYARAAGLYVDAHGTGVAAGLYFNGRLAMSDDLPLPVQYQVLTTPNSDLKT 2
Db	281 LGTAKMLFNTGLFLGLDVPTGGLGVANAGLYFNGHLLAMSDDLPLPVYHVTOTSDLET 3
Qy	281 VGRFDQCLESTMIAHPKVDESEGEFALSVDVSKPVLXWFRFSPDGTSPPDVEIQLD 3
Db	341 SGYDFDQLKSTMIAHPKIDDETFEFLASDVSVSKPVLXWFRFSDGEPSDVEIPD 4
Qy	341 QPMHMHDFAITENVVYVDPQQVYKLFPEMIRGSPVYYDKNKVAFRGFLDKYAEDDSNN 4
Db	401 QPMIHDAITENVVYDQVQYERLPMKMRGSPVYYDEKKRSRFGFLNKAKDASIQ 4
Qy	401 WIDAPPDCFCFHLMNAWEPEPETDEVVIGSCMTTPDSIFNEDENLKSVLSETRLNLKTE 4
Db	461 WIEVDPDCFCFHLMNSWEPEPETDEVVIGSCMTPEPDSEHETLQSVLSETRLNLKTE 5
Qy	461 STRRPITISNEDQOYNLEAGMVNRMLGRTKFAYLALAEPWPKVSGFAKVDLTGEVKKH 5
Db	521 STRREVIS--EQYNLEAGMVNRMLGRKTRAYLATLPEWPKVSGFAKVDLSTGEIRKY 5
Qy	521 LYGDNRYGGFPFLPGEGFEDSYLGFEGVDEYKLVLSRVLKSLQIYVAVKLSRV 5

Qy 417 BEPETDEVVIGSCMTPPDSTINESDENIKSVLSEIRLNKTEGESTRPIISNEDQQVN 476
 Db 442 EEPESSEVVVYSCMTPPDSTINESDENIKSVLSEIRLNKTEGESTRPIISNEDQQVN 476
 Rx 477 EAGMYRNMLGRKTREPAYLAIEPWPKVSGFAKVLDTGEKYKHLGYDNYGEPLFLPG 536
 Ra White O., Alonso J., Aclaf H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewart K.,
 RA Dunn P., Etuq P., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Gill J.E., Jenkins J., Johnson-Ropson C., Khan S., Khaykin E.,
 RA Hunter J.L., Koo H.I., Krementskaya I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luos R., Maiti R., Marziali A.,
 RA Miltsccher J., Miranda W., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shian P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambang G., Toriumi M.J., Town C.D.,
 RA Utterback T., Vijn Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Ventier J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the Plant Arabidopsis
 thaliana.";
 RL "816-82(2000)."
 DR EMBL; AC04176; AGI0855; 1; -.
 DR InterPro; IPR00494; RPE65.
 DR Pfam; PF03055; RPE65; 1.
 KW Dioxygenase.

SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 MEDLINE=21:016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Fedderspiel N.A., Kaul S.,
 RA White O., Alonso J., Aclaf H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewart K.,
 RA Dunn P., Etuq P., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Gill J.E., Jenkins J., Johnson-Ropson C., Khan S., Khaykin E.,
 RA Hunter J.L., Koo H.I., Krementskaya I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luos R., Maiti R., Marziali A.,
 RA Miltsccher J., Miranda W., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shian P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambang G., Toriumi M.J., Town C.D.,
 RA Utterback T., Vijn Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Ventier J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the Plant Arabidopsis
 thaliana.";

DR EMBL; AC04176; AGI0855; 1; -.
 DR InterPro; IPR00494; RPE65.
 DR Pfam; PF03055; RPE65; 1.
 KW Dioxygenase.

SEQUENCE 589 AA; 65336 MW; 6A79B2BF7AA8179 CRC64;

Query Match 64.9%; Score 2044.5; DB 10; Length 589;
 Best Local Similarity 68.1%; Pred. No. 4.2e-151; Mismatches 69; Indels 13; Gaps 4;

Matches 390; Conservative 69; Mismatches 101; Indels 13; Gaps 4;

Qy 27 SQSSDISYCSSSLPMASRVTKLNVSSALHPTPAHFPKOSSNSPAIVVKPKAKSNTKOM 86
 Db 27 SPSSPSVFTNTPKPR ---RKLANSVSDPPLNFNPVSPNPVSPNPI ---PEKDTSRW 76

Qy 87 NLQRRAAALDAAGFELVSHKELPKPTADPSVQAGNFAFPVNPYQVRNNPVPVGKLP 146
 Db 77 NPLQRAASALDFATALLRRRSKPLPKTVDPRHQISGNYAPPEQSYYKSSLVSDGKLP 136;

Qy 147 DSIKSVYVRNGANPLHEPYTGHHFDGDMVHAYKFEHGSASYACRFQTNRFYQERQLG 206

Qy	417 BEPETDEVVIGSCMTPPDSTINESDENIKSVLSEIRLNKTEGESTRPIISNEDQQVN 476	Db	137 DCIDGYYLRLNRGANCPLPEPVSSHHLEFGDMVHAYKITEERLYQEVKQLG 196
Db	442 EEPESSEVVVYSCMTPPDSTINESDENIKSVLSEIRLNKTEGESTRPIISNEDQQVN 476	Qy	207 RPVFPKAIGELGHGTGTRALMFLYARAAGTYPDAHGTVGANAGLVYDASYSCHETETERLYQEVKQLG 196
Qy	477 EAGMYRNMLGRKTREPAYLAIEPWPKVSGFAKVLDTGEKYKHLGYDNYGEPLFLPG 536	Db	197 SPIFFKAIGELGHGTGTRALMFLYARAAGTYPDAHGTVGANAGLVYDASYSCHETETERLYQEVKQLG 196
Db	501 EAGMYRNRLGRKTREPAYLAIEPWPKVSGFAKVLDTGEYNFVYGERFGEPEYFIPR 560	Qy	267 PYQVOTPGNGDLKTVGRDFGQLESTMIAHPKVDPESGLFALSYDVSKPALKYFRS 326
Qy	537 B---GGEDEGYLICCFVHDEXTWKSELQTYNAVSLEVEATYKLPSPRVYGFHGTIGADD 593	Db	257 PYQVRTDINGGLETIGRDFGQLSMAMIAHPKVDPESGLFALSYDVSKPALKYFRS 316
Db	561 EPSTSGREDDGYYVVFHDEKTSRSELLINAMNNRLEASVMLPSKRVYGFHGTFISSRO 620	Qy	327 PGTKSPDVEIQLDQPTMMHDFATENFVYVPDQQVVFKEPEMIGGSPVYDANKVAFR 386
Qy	594 LAQK 597	Db	317 PEGKSPDVEIPLASPTMMHDFATENFVYVPDQQVVFKEPEMIGGSPVYDANKVAFR 376
Db	621 LAQK 624	Qy	387 GILDQYAEDSSNIKWLADPDCFCPHLNWAAPEPDEVVIGSCMTPPDLSLFNESENLK 446
RESULT 9	Q9C6Z1 ID Q9C6Z1_ PRELIMINARY; PRT; 589 AA.	Db	377 GILPRAKDASEMWWVSEPETECFHLWNAAWESPDEVVIGSCMTPDLSIFNECDEQLN 436
AC	Q9C6Z1_	Qy	447 SVLSEIRLNLRKTGSTRPILSNEDQQVNLEAGMYNRMNLGRKTKFAYLALAEPWPKVSG 506
DT	01-JUN-2001 (TrEMBLrel. 17, Created)	Db	437 SVLSEIRLNLRKTGSTRTIPLG-SVQMVLEAGMYNRMNLGRKTAYAIAEPWPKVSG 495
DT	01-JUN-2002 (TrEMBLrel. 17, Last sequence update)	Qy	507 FAKVDTITGEVKKHYDNYGGEPFLPLPG -EGGEDEGYILCFVHDEKTKWSELQIN 564
DB	9-cis-epoxycarotenoid dioxygenase, putative.	Db	496 FAKVDSLSTGEVKNHFYGGKRYGGEPFLPLRGLESDEGGYMSFVHDESWESLHTVN 555
GN	T2H7.10	Qy	565 AVSLEVEATKLPSPRVYGPFGHGTFLGADLAKQ 597
OS	Arabidopsis thaliana (Mouse-ear cress).	Db	556 AVTLELEATVKLPSKRVYGPFGHGTFLGTVNSADMNLQ 588
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
RN	[1]		
RP	SEQUENCE FROM N.A.	RESULT 10	
RC	NCBI TAXID=3702;	ID 049505 PRELIMINARY; PRT; 583 AA.	
RX	049505; AC 049505; DT 01-JUN-1998 (TrEMBLrel. 06, Created)	AC 049505; DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)	
RA	Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewart K., Dunn P., Etuq P., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L., Gill J.E., Jenkins J., Johnson-Ropson C., Khan S., Khaykin E., Kim C.J., Koo H.I., Krementskaya I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luos R., Maiti R., Marziali A., Miltsccher J., Miranda W., Nguyen M., Nieman W.C., Osborne B.I., Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shian P., Southwick A.M., Sun H., Tallon L.J., Tambang G., Toriumi M.J., Town C.D., Utterback T., Vijn Aken S., Vaysberg M., Vysotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Ventier J.C., Davis R.W.; RT "Sequence and analysis of chromosome 1 of the Plant Arabidopsis thaliana.";	RA 049505; DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RN	[1]	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RP	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	
RA	Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A., Banchroft I., Newes H.W., Mayer K., Schueler C.; Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.	RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A., Banchroft I., Newes H.W., Mayer K., Schueler C.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RN	[2]	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RP	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	
RA	Murphy G., Ridley P., Hudson S., Newes H.W., Lemcke K., Mayer K.F.X.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	RA Murphy G., Ridley P., Hudson S., Newes H.W., Lemcke K., Mayer K.F.X.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RN	[3]	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RP	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	
RA	Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A., Newes H.W., Lemcke K., Mayer K.F.X.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A., Newes H.W., Lemcke K., Mayer K.F.X.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RN	[4]	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RP	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	RA 049505; DT 02812.10 OR AN4G18350; DE F2812.10 OR AN4G18350; OS Arabidopsis thaliana (Mouse-ear cress). RA 049505; DT 02812.10 OR AN4G18350; OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosid II; Brassicales; Brassicaceae; Arabidopsis. NCBI TAXID=3702;	
RA	RA 049505; DT 02812.10 OR AN4G		

Qy	16	GGNHTDPPPLSSQSSDLSY -- -CSSLPMASRYTRKLNVSSALHTPPALHEPKOSSNSPATV	73	Db	9	KTFTHSHSLHSSSPVPHLPKLLTTPLHEKSQRELGLIL ---OEPNRAKWNFQRAAA
Db	11	GGIKTWR -- -QAQIDLGFRPIKRQPKVKTQIVDTVE-LTKKROLEPTPRITATPP -- 62	Qy	95	AALDAEAFGLYSH - -EKLHPDLPKTAQPSVQTAGNAPVNQPYVRNLVVQKLPDSIKGV	
Db	74	VPKAKESNTIKOMNLFORAAAALDAEGFLYSHEKHLPLPKTADPSVQLAGNAPVNEQ	133	Db	65	VALDTVEDSFISGVLERRHLPLKTSIDPAVOISGNAPAYDEHPYQHHLBVSGRTRCLGV
Db	63	QHNPRLNLIQFOKAATAIDAEARLISHEQDPLPKTADPVRQIAGNYSPVBS	116	Qy	153	YVRNGANPLHPEVTGHHFEDGDMWHAVKFEHGSASYACRFQTNRFVQYEROLGRPVFKP
Db	117	SVRNLIVLTGTTIPDCIDGVYFRGANMFPETAGHDFGDGMWHAVKTNGASYACRF	193	Db	125	YLRNGANPLPVGHHFEDGDMWHSLSRGTASTACRFETHRLVQERAIGRPVFKP
Db	117	YARAAAGYDPAHGNGVANGLYVNRLLAMSEADLPYQVQI	272	Qy	213	ALGEFLHGTGIAIRMLFYARAAAGYDPAHGNGVANGLYVNRLLAMSEADLPYQVQI
Db	117	YARAAAGYDPAHGNGVANGLYVNRLLAMSEADLPYQVQI	272	Db	185	AGEFHGHSGARLLFYARATGFLYDGSSTGVAPAGLVENRHLJAMSEADLPYQVQI
Qy	194	TQTNREVQERQLGRPVFPKAIGELHGTGIAIRMLFYARAAAGYDPAHGNGVANGLYV	253	Qy	273	TPNGDLKTVGRFDQDQLESTMIAHKVDPGSGELFAALSVDVSKPVLKFRTSPDGTKS
Db	177	TKTERLVQEKRQKLRGRPVFPKAIGELHGTGIAIRMLFYARGLGLINNQNGVANGLYV	236	Db	245	TSSGDLTETVGRFDQEMGQLNSAMIAHKLDPASGEFLAFLSYNTVKPKLKFRTSPDGKKS
Qy	254	FNGRLAMSEDDELPLPVQVQITPQGDLKTVGRFDQGDESTMIAHKVDPGSGELFALSYD	313	Qy	333	PDVETQDQPTIMMHDFAITENFVVPDQVYKPEMTRGGSPPVYDKNKYARFGFLDKY
Db	237	FNRNLAMSEDDELPLPVQKLTQTDQKJXSAMAHPKLDPVTELHALSYD	296	Db	305	PDVETQDQPTIMMHDFAITENFVVPDQVYKPEMTRGGSPPVYDKNKYARFGFLDKY
Qy	314	VVSKPKLYFRFSPDGKSPKPEEPIPLPQVYKPEMTRGGSPPVYDKNKYARFGFLDKY	373	Qy	393	AEDSNKIKWIDADPCFCFHWNAAWEPEPETDEVVIGSCMTTPDSTFNESENULKSYLSEI
Db	297	VVKPKLYFRFSPDGKSPKPEEPIPLPQVYKPEMTRGGSPPVYDKNKYARFGFLDKY	356	Db	365	AADSNGLRWDADPCFCFHWTAWEEPETDQVWVIGSCMTTPDSTFNESENULKSYLSEI
Qy	374	SPVYDKNKYARFGFLDKYADEDDSSNIKWKIDADPCFCFHWNAAWEPEPETDEVVIGSCMTP	433	Qy	453	RNLNTKGESTERRPISNEDQVNLNEGMVNRMNLGRKTKFAYLAEPWPKVSGAKVYL
Db	357	SPVVFGEKEVRLGIMPKDATEASQIWIWNSPETCFCHWNAAWEPETEIVVIGSCMSP	416	Db	425	RNLNTKGLOSSREI - DPSRHLNLVYGMVNRMNLGRGPGVSLSSHCRWPVKGAKVYL
Qy	434	PDSFINESDPLSKVSLSEIRLNLKGESTRPPIISNEDQYRQLEAGMVNRNMIGRKTKF	493	Qy	513	TTGENKKHLIGDNRYGGEPFLPQEG - GEEDEGYTLFCVHDEKTKWSLQITVNAVSLEVE
Db	417	ADSTNERDEELRSVLSLEIRLNLRKTKTSSLVNEY - VNLIGMWNRNLGRKTKF	474	Db	483	STGETKFLFIGEQCYGGEPFVSRDPVAPEDDGYVLSIFMHDDEKTAVERSLLVNTWOLE
Qy	494	YLALAEPPWPKVSGFKAVYKDLTGEVKHYDGNRYGGEPFLPQEGG -- - FEDEGYTLFC	549	Qy	572	ATVKPLSPRVYPYGFHGTFIGADDLAKQ
Db	475	FLAIAYPPWPKVSGFKAVYKDLTGEVKHYDGNRYGGEPFLPQEGG -- - FEDEGYTLFC	534	Db	543	ASVKLPLSPRVYPYGFHGTFISSKDLAQ
Qy	550	YHDETKWSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	598	RESULT 12		
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	ID	024592	PRELIMINARY;
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	AC	024592;	PRELIMINARY;
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DT	024592;	PRT;
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DT	01-JAN-1998	PRT;
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DT	01-JAN-1998	(TREMBLrel. 05, Created)
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DE	01-JAN-2001	(TREMBLrel. 19, Last annotation update)
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	GN	Viviparous-14 .	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	OS	Zea mays (Maize).	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade.	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	OC	Panicoideae; Andropogoneae; Zea.	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	OX	NCBI_TaxID=4577;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RN	[1]	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RP	SEQUENCE FROM N.A.	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RC	STRAIN=N22;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RA	SEQDBNAME=N.A.	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RA	MEDLINE=98004553; PubMed=9342392;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RA	Tan B.C., Schwartz S.H., Zeevaart J.A., McCarty D. R.; Zeevaart J.A., McCarty D. R.;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RA	"Genetic control of abscisic acid biosynthesis in maize." Proc. Natl. Acad. Sci. U.S.A. 94:12235-12240 (1997).	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	RL	EMBL; U09553; AAB6281; 1;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DR	InterPro; IPR004294; RPE65.	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DR	Pfam; PF03055; RPE65; 1;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	SQ	SEQUENCE 604 AA; 65587 MW; CBCDC4DB85F61BE CRC64;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	SQ	Query Match 61.3%; Score 1930; DB 10; Length 604;	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DB	Best Local Similarity 63.3%; Pred. No. 3.9e-142; Mismatches 373; Conservative 73; Indels 22; Gaps	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	Qy	22 PPLSSQSSDLSYCSLPMASRYTRKLNVSSALHTP -- -PALHEPKQSSNSPAIVY --- 74	
Db	535	YHDETKTSLQIYNAVSLEATEVYKLPSPRVYPYGFHGTFIGADDLAKQV	583	DB	17 PAISRAS --- -NSVRSRAPSVPYVAECLOPFHPVADIPAPSRSRKPAAVPGCHA 71	

QY	75 -KPKAKESNTKOMNLQRAAAALDA-AEGELVS-HEKLHLPLKTADEPSVQTAGNEAPVN 131	QY	24 LSSQQSSDLSYCSSLPMASRVTRKLNVSS-ALHTPPAIHFPKQSSNSPATIVKPKAKESN 82
Db	72 APRKAKEGGKQNLQRAAAALDAEFGEGVANPLHEPVTGHFFDGMVHAYKAEFGSA-SYA 131	Db	9 LLPTKTKPSRSHLPOPANANSLRILINPEKIPTLPLDTSPVPSP-----VKLKPTYPN 62
Qy	132 EOPVRNLPPVYKLPOSIGKVVRNGANPLHEPVTGHFFDGMVHAYKAEFGSA-SYA 190	Qy	83 TKOMNLQFRAAAAALDAE-GFLVSHBKRLPLKTADEPSVQTAGNFAPVN EQPVRENLPV 141
Db	132 EPPVHELPVSRRIPPLDGYARANGANPCPDPVAGHHLFDDGMVHAYKAEFGSA-SYA 191	Db	63 --LNLLQKLAATMLDKIESSTIVIPMEQNRPLPKPDPAVQLSGNFAVN EC PVQNGLEV 119
Qy	191 CREFTQTNRFVQERQLGRPVFPKAIGELHGHTGJLARIUMLFYARAAGA GTIVDPAHGTGYNAG 250	Qy	142 VKGKLPDSIKGVYVVRNGANPLHEPVTGHFFDGMVHAYK-FERGSA-SYACRFTOTNRF 199
Db	192 CREFETARLRLQERAIAGRPPVFPKAIGELHGHTGJLARIUMLFYARAAGA GTIVDPAHGTGYNAG 251	Db	120 VGGIPSCKIGVYIRNGAMPFPLA GHLFEDDMMAVHSIGFDN-QVSYCRYKTINRL 178
Qy	251 LYTFNGRLAMSEDDLPLQYQVQITPNGDLKTVGREDFGQLESTMIAHPKVDPDESGELFAL 310	Qy	200 VQEROLGRPVFEKAIGELHGHTGJLARIUMLFYARAAGI VDPAHGTGYNAGLVYFNGRL 259
Db	252 LYTFNGRLAMSEDDLHYVRAADDGLLETYGRYDPSQLGCA MTAHPKLPATGELHAL 311	Db	179 VOETALRSVPEKPKIGDQHGSGLARALTA RAGIGLVDSTRGMVANAGYVFENGRL 238
Qy	311 SYDVVSKPVLKYYFRPSDGTKSPDVEIQLODFTMMHDFAITENFVVVPPDQGVVFKEPEMI 370	Qy	260 AMSEDDPLPYQVQIITPNQDLKTYGRFEDFGQLESTMIAHPKVDPDESGELFALSYDVSKPY 319
Db	312 SYDVVKKPVLKYYFRPSDGTKSDDVIPLEQTMIDHDEFAITENFVVVPPDQHVVFKEQML 371	Db	239 AMSEDDPLPYQVQIKDQGDLLETIGRFQFDQIDSSVLAHPKVDATGD LHTLSYNVLKPH 298
Qy	371 RGGSPPVYYDKNVA REGILDYKAEDSSNIKWTADPDCFCFLWNAWEPEPDEVVVGSC 430	Qy	320 LKYFREPSDGTKSPDVEIQLODFTMMHDFAITENFVVVPPDQGVVFKEPEMI 379
Db	372 RGGSPPVYYDKNVA REGILDYKAEDSSNIKWTADPDCFCFLWNAWEPEPDEVVVGSC 431	Db	299 LRYLKENTGKTRDVEITLPEPTMAHDFAITENFVVVPPDQMVFKLSEMTRGGSVIVY 358
Qy	431 MTPPDSIFNESDENLKSVLSEIRLNLTGTESTRPLISNEDQVNLEAGMVNRMIGRKT 490	Qy	380 XKEKMRAGVFLSKQDLTGSDINWVDVPCFCFLWNAWEPEPDEVVVGSCMTIPD 435
Db	432 MTPPDSIFNESDENLKSVLSEIRLNLTGTESTRPLISNEDQVNLEAGMVNRMIGRKT 490	Db	359 KEKMRAGVFLSKQDLTGSDINWVDVPCFCFLWNAWEPEPDEVVVGSCMTIPD 417
Qy	491 KFAYLALAEPNPKVSGFAKYLDTGEVKKHLYGDNRYGGPELJLPGEGGE----EDEGY 545	Qy	436 SIFNESDENLKSVLSEIRLNLTGTESTRPLISNEDQVNLEAGMVNRMIGRKT 495
Db	491 KFAYLALAEPNPKVSGFAKEDISTGE LTGEVKKHLYGDNRYGGPELJLPGEGGE----EDEGY 545	Db	418 TIFSESEPTVRLSEIRLNLTGTESIRKRVITVG---VNLEAGHINRSTYGRKSOFVYI 473
Qy	546 ILCFYHDEKTRKSELQIVNAYSLLEVATVKLPSRVYGFHGTFIGADDLAKQ 597	Qy	496 ALAEPNKVKVSFAKVPLDTGEVKKHLYGDNRYGGPELJLPGEGGE----EDEGY 555
Db	551 VLFVHDERAGTSELVNVNAADIREATVOLPSRVYGFHGTFIGADDLAKQ 602	Db	474 AIADPNPKCSGIAKVDIQNGTVSEFNYGPSFGEFCPVPERGEDEGKYMGFVRDEK 533
QY	556 WKS ELQIVNAYSLLEVATVKLPSRVYGFHGTFIGADDLAKQ 598	Qy	534 DESEFVVVDA+DMKQVAAVRUPERPVYFGFHSTFVSENQLEQV 576
Db	534 DESEFVVVDA+DMKQVAAVRUPERPVYFGFHSTFVSENQLEQV 576	Db	534 DESEFVVVDA+DMKQVAAVRUPERPVYFGFHSTFVSENQLEQV 576
<hr/>			
RESULT 13			
ID	Q9LRM7 PRELIMINARY;	PRT;	577 AA.
AC	Q9LRM7 PRELIMINARY;	PRT;	595 AA.
DT	01-OCT-2000 (TRMBLrel. 15, Created)	049675 PRELIMINARY;	PRT;
DT	01-OCT-2000 (TRMBLrel. 15, Last sequence update)	049675 PRELIMINARY;	PRT;
DT	01-JUN-2002 (TRMBLrel. 21, Last annotation update)	049675 PRELIMINARY;	PRT;
OS	Arabidopsis thaliana (Mouse-ear cress).	AC	049675;
OC	Eukaryota; Viridiplantae; Streptophytina; Embryophytina;	DT	01-JUN-1998 (TRMBLrel. 06, Created)
OC	Spermatophytina; Magnoliophytina; eudicots; Rosidae;	DT	01-JUN-1998 (TRMBLrel. 06, Last sequence update)
OC	eurosidia II; Brassicales; Brassicaceae; Arabidopsis.	DT	01-DEC-2001 (TRMBLrel. 19, Last annotation update)
NCBI_TAXID=3702;		DE	NEOXANTHIN cleavage enzyme-like protein (At4G19170/T1B16_140 OR At4G19170).
RN	SEQUENCE FROM N.A.	GN	Arabidopsis thaliana (Mouse-ear cress).
RC	STRAIN-COLUMBIA;	OS	Eukaryota; Viridiplantae; Streptophytina; Embryophytina;
RA		OC	Spermatophytina; Magnoliophytina; eudicots; Rosidae;
RA	Nakamura Y., Kaneko T., Asamizu E., Tabata S.;	OC	eurosidia II; Brassicales; Brassicaceae; Arabidopsis.
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.	OX	NCBI_TaxID=3702;
RN	[2]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN-COLUMBIA;	RA	Bevan M., Benes V., Rechmann S., Borkova D., Ansorge W., Bancroft I.,
RX	MEDLINE=20277480; PubMed=10819329;	RA	Newes H.W., Mayer K., Schueler C.;
RA	Nakamura Y.;	RA	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty P1 and TAC clones.";	RN	[2]
RT	RT	RN	SEQUENCE FROM N.A.
RL	DNA Res. 7:131-135(2000).	RA	Benes V., Rechmann S., Borkova D., Ansorge W., Newes H.W., Lemcke K.,
DR	EMBL: AB028621; BAB01363.1; -.	RA	Mayer K.F.X.;
DR	Introp; IPR004294; RPE65.	RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR	Pfam: PF03055; RPE65; 1.	RN	[3]
KW	Dioxxygenase.	RN	SEQUENCE FROM N.A.
SQ	SEQUENCE 577 AA; 63821 MW; 515F0A1B321B4F1 CRC64;	RA	EU Arabidopsis sequencing project;
Query Match	52.8%; Score 1663.5%; DB 10; Length 577;	RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
Best Local Similarity	55.4%; Pred. No. 2.4e-121;	RL	[4]
Matches	Conservative 100; Mismatches 137; Indels 23; Gaps 9;	RA	SEQUENCE FROM N.A.
Cheuk R., Chen H., Kim C.J., Koessema E., Meyers M.C., Banh J.,	RA	RP	

